

## BIOSYNTHETIC BINDING PROTEINS FOR IMMUNO-TARGETING

The U.S. Government may have certain rights in the invention described herein, by virtue of National Institutes of Health Grant No. UO1 CA51880.

5

### **Cross Reference to Related Applications**

This application is a continuation-in-part of U.S. Serial No. 09/558,741, filed April 26, 2000, which is a continuation-in-part of U.S. Serial No. 07/831,967 filed February 6, 1992. This application is also a continuation-in-part of U.S. Serial No. 08/575,724 filed 10 Dec. 18, 1995, now US Patent No. 6,207,804, which is a continuation of U.S. Serial No. 08/139,901, filed Oct. 19, 1993, now U.S. Pat. No. 5,476,786, which is a continuation of U.S. Serial No. 07/850,228, filed Mar. 12, 1992, now abandoned, which is a continuation of U.S. Serial No. 07/213,671, filed Jun. 30, 1988, now U.S. Pat. No. 5,132,405, which is a continuation of U.S. Serial No. 07/052,800, filed May 21, 1987, now abandoned. The 15 specifications of each of the foregoing are hereby incorporated by reference in their entirety.

### **Background of the Invention**

The development of murine monoclonal antibodies and their proteolytic Fab fragments has raised interest in their utility as diagnostic and therapeutic reagents for in vivo imaging and drug targeting. However, successful in vivo targeting of radionuclides, 20 drugs or toxins using 150 kD intact antibodies or their 50 kD Fab fragments (an antibody fragment consisting of one light chain and approximately half of the heavy chain held together by a single disulfide bond) have been restricted by the limited penetration of these molecules from the vasculature into the tissues of interest, and by their slow clearance 25 rates in vivo, which for IgG leads to behavior that requires several days to clear the background enough for imaging to be possible. Other disadvantages of the intact antibodies or their Fab fragments include: their immunogenicity when prepared from different species, their non-specific binding to many normal tissues and organs, and the 30 fact that they contain multiple proteolytic cleavage sites which result in their degradation during their circulation in vivo.

Although Fv fragments, which consist of one  $V_H$  and one  $V_L$  domain held together by noncovalent interactions, form the minimal region of an antibody that contains a

complete antigen combining site, dissociation of the  $V_H$  and  $V_L$  domains *in vivo* can preclude their use as therapeutic or imaging agents. Although Moore *et al.*, (U.S. Pat. No. 4,642,334) and Glockshuber *et al.*, (1990, *Biochem.* 29, 1362-1367) disclose attempts to stabilize these Fv fragments with engineered intermolecular disulfide bonds, monovalent 5 50 kD Fab and Fab' fragments have, until recently, been the smallest proteins available for effective immunotargeting.

Recently, single-chain Fv (sFv) polypeptide chains of about 27 kD have been developed containing covalently linked  $V_H$ - $V_L$  polypeptides. The  $V_H$  - and  $V_L$  -domains are connected by a polypeptide linker. The resulting sFv polypeptide chains are also referred to in the art as biosynthetic antibody binding sites or BABS and preferably are 10 encoded by a single DNA sequence. For a detailed description of these biosynthetic polypeptide chains see for example, Huston *et al.*, 1988, *Proc. Nat. Acad. Sci. USA* 85: 5879-5883 or U.S. Pat. Nos. 5,091,513 and 5,132,405, all of which are hereby incorporated by reference. The sFv polypeptide chains provide attractive alternatives to 15 intact immunoglobulins and Fab fragments due to their small size and their stability at concentrations that typically promote dissociation of natural Fv fragments. U.S. Pat. Nos. 5,091,513 and 5,132,405; Huston *et al.*, ((1991) *Methods in Enzymology* 203: 46-88; Huston *et al* (1993) *Int. Rev. Immunol.* 10: 195-217) disclose the utility of sFv polypeptides, as well as single chain constructs synthesized from single DNA sequences, 20 which may further comprise ancillary effector proteins, such as a second sFv or a cytotoxic agent.

Pack *et al.* ((1992) *Biochem* 31: 1579-1584) disclose the construction of "mini-antibodies". The mini-antibodies are sFv polypeptide chains which also include an "oligomerization domain" at their C-termini, separated from the sFv by a hinge region. 25 The oligomerization domains comprise self-associating  $\alpha$ -helices, for example, leucine zippers, that can be further stabilized by additional disulfide bonds. The domains are designed to be compatible with vectorial folding across a membrane, a process thought to facilitate *in vivo* folding of the polypeptide into a functional binding protein.

PCT application PCT/US92/09965, published Jun. 10, 1993 also discloses the 30 construction of bivalent sFv constructs, including crosslinked dimers. However, the pharmacokinetic properties of these constructs or those disclosed by Pack *et al.* are not measured *in vivo*.

PCT application PCT/US92/07986, published Apr. 1, 1993, discloses bifunctional (Fab')<sub>2</sub> molecules composed of two Fab' monomers linked through cysteine residues located at the C-terminus of the first constant domain of each heavy chain. PCT application PCT/US92/10140, published Jun. 10, 1993, also discloses bifunctional (Fab')<sub>2</sub> dimers which, in addition to the cysteine residues located in the hinge region, also contain C-terminal leucine zipper domains that further stabilize the (Fab')<sub>2</sub> dimers. In both cases, the resulting (Fab')<sub>2</sub> dimers ( $\geq 100$  kD in size), although smaller than intact immunoglobulins, are significantly larger than sFv polypeptides and are anticipated to have slower tissue biodistribution and clearance rates following in vivo administration.

10 Cumber et al. disclose the generation of (Fv-Cys)<sub>2</sub> heterodimers by chemically crosslinking two V<sub>H</sub> -cys domains together (Cumber et al., 1992, J. Immunology **149B**: 120-126). Although the crosslinked V<sub>H</sub> chains appear to be stable, dissociation of the V<sub>L</sub> polypeptides from each Fv reduces the pharmacological value of these constructs in vivo.

15 It is an object of the instant invention to provide biosynthetic constructs having enhanced pharmacokinetic properties as in vivo targeting agents. In particular, it is an object of this invention to provide biocompatible constructs having accelerated in vivo biodistribution and body clearance rates in comparison to whole antibodies or antibody fragments. It is another object of the invention to provide biosynthetic constructs having enhanced avidity in vivo, including enhanced target tissue specificity and target tissue 20 retention. Yet another object is to provide dimeric biosynthetic constructs having improved tissue imaging and drug targeting properties in vivo. Still another object is to provide diagnostic and therapeutic formulations comprising these constructs. Such 25 formulations have particular utility in the diagnosis and treatment of malignancies. Still another object is to provide constructs having enhanced pharmacokinetic properties as in vivo targeting agents, particularly as in vivo imaging agents, for ovarian and breast tumor tissue.

These and other objects and features of the invention will be apparent from the description, figures and claims which follow.

30 **Summary of the Invention**

A class of novel biosynthetic polypeptides has now been designed and engineered which comprise biosynthetic antibody binding sites, that is, "BABS" or chimeric polypeptides defining structure capable of selective antigen recognition and preferential antigen binding.

In its broadest aspects, this invention features polypeptides comprising biosynthetic antibody binding sites, DNA encoding these polypeptides prepared by recombinant DNA techniques, vectors comprising these DNAs, and methods for the production of these polypeptides.

5 In one aspect, the invention is based on the observation that three subregions of the variable domain of each of the heavy and light chains of native immunoglobulin molecules collectively are responsible for antigen recognition and binding. Each of these subregions, called herein "complementarity determining regions" or CDRs, consists of one of the hypervariable regions or loops and of selected amino acids or amino acid sequences disposed in the framework regions which flank that particular hypervariable region. It has now been discovered that framework regions from diverse species are effective to maintain CDRs from diverse other species in proper conformation so as to achieve true immunochemical binding properties in a biosynthetic protein. Thus, BABS produced in accordance with the invention comprise biosynthetically produced novel sequences of 10 amino acids defining polypeptides designed to bind with a preselected antigenic material. The structure of these synthetic polypeptides is unlike that of naturally occurring antibodies, fragments thereof, or known synthetic polypeptides or "chimeric antibodies" in that the regions of the BABS responsible for specificity and affinity of binding, (analogous to native antibody variable regions) are themselves chimeric, e.g., comprise amino acid 15 sequences homologous to portions of at least two different antibody molecules.

20

The invention thus provides a chimeric polypeptide defining a region capable of selective antigen binding and recognition. This chimeric polypeptide comprises amino acid sequences homologous to portions of the CDRs of the variable domain of one immunoglobulin light or heavy chain, and other sequences homologous to the framework 25 regions, or FRs, of the variable domain of a second, different immunoglobulin light or heavy chain. Polypeptides so constructed bind a specific preselected antigen determined by the CDRs. Preferably, the chimeric polypeptides comprise an amino acid sequence homologous to at least a portion of the variable regions of a mammalian immunoglobulin, such as those of mouse, rat, or human origin. In one preferred embodiment, the 30 biosynthetic antibody binding site comprises FRs homologous with a portion of the FRs of a human immunoglobulin and CDRs homologous with CDRs from a mouse immunoglobulin. This type of chimeric polypeptide displays the antigen binding specificity of the mouse immunoglobulin, while its human framework minimizes human immune reactions. In addition, the chimeric polypeptide may comprise other amino acid

sequences. It may comprise, for example, a sequence homologous to a portion of the constant domain of an immunoglobulin, but preferably is free of constant regions (other than FRs).

The invention also provides a single chain composite polypeptide having antigen binding abilities, and comprising a pair of amino acid sequences homologous or analogous respectively to the variable regions of an immunoglobulin light and heavy chain, (linked  $V_H - V_L$  or single chain Fv). Both  $V_H$  and  $V_L$  may copy natural monoclonal sequences, or one or both of the chains may comprise a CDR-FR construct of the type described above. The separate polypeptides analogous to the variable regions of the light and heavy chains are held together by a polypeptide linker.

This type of chimeric polypeptide is thus a single chain composite polypeptide comprising a complete antibody binding site. This single chain composite polypeptide has a structure patterned after tandem  $V_H$  and  $V_L$  domains, but with the carboxyl terminal of one attached through an amino acid sequence to the amino terminal of the other. It thus comprises an amino acid sequence which is homologous to a portion of the variable region of an immunoglobulin heavy chain ( $V_H$ ) peptide bonded to a second amino acid sequence which is homologous to a portion of the variable region of an immunoglobulin light chain ( $V_L$ ). The linking amino acid sequence may or may not itself be antigenic or biologically active. In addition, either the amino or carboxyl terminal ends of these chimeric, single chain Fvs may be attached to an amino acid sequence which itself is bioactive to produce a bifunctional or multifunctional protein. For example, the synthetic Fv may include a leader or trailer sequence defining a polypeptide having enzymatic activity, independent affinity for an antigen different from the antigen to which the chimeric Fv is directed, or having other functions such as to provide a convenient site of attachment for a radioactive atom, or simply to enhance expression in prokaryotic host cells or yeasts.

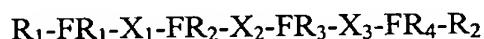
Such tandem arrangement of  $V_H$  and  $V_L$  polypeptides can increase the stability of the antigen binding site and facilitate its coupling to proteins utilized in drug targeting and moieties useful in imaging. The therapeutic use of such chimeric Fvs provide a number of advantages over larger fragments or complete antibody molecules: they are often quite stable and less immunogenic; they can penetrate body tissues more rapidly for purposes of imaging or drug delivery because of their smaller size; and they can facilitate accelerated clearance of targeted isotopes or drugs.

Other embodiments of the invention comprise multifunctional polypeptides consisting of one or more single chain Fvs either linked  $V_H$  and  $V_L$  dimers, individual  $V_L$

or  $V_H$ , or any of the foregoing comprising CDRs and FRs from different or the same immunoglobulins, linked to a second functional protein domain such as, for example, a toxin, enzyme, or site of attachment to an immobilization matrix. Yet another embodiment is a polypeptide comprising several identical or non-identical BABS which recognize a group of antigenic determinants that are periodic or closely spaced in their normal environment, e.g., on a cell surface. This arrangement confers greatly augmented affinity and/or specificity on the BABS-containing protein analogous to, for example, the way IgM (containing 10 Fabs) binds to the surfaces of certain cells.

In other aspects, the invention provides DNA sequences encoding chimeric polypeptides of the type described above, vectors including such sequences, and methods employing the DNAs and vectors for producing the polypeptides.

A novel method of producing BABS involves the construction of a DNA containing three polynucleotide sequences ( $X_1$ ,  $X_2$  and  $X_3$ ). Each of the sequences contain restriction sites proximal its 3' and 5' ends, and each is flanked by polynucleotide sequences ( $FR_1$ ,  $FR_2$ ,  $FR_3$  and  $FR_4$ ) encoding selected framework region (FR) amino acid sequences homologous to a portion of the variable domain of an immunoglobulin. This DNA has the structure:



where  $R_1$  is a 5' phosphate group or polynucleotide sequence and  $R_2$  is a 3' hydroxyl group or polynucleotide sequence. The  $X$  polynucleotide sequences may be selectively excised using restriction enzymes and replaced by other DNA sequences encoding the CDR amino acid sequences of a variable domain of a selected immunoglobulin. This type of DNA sequence may encode at least part of the variable region of either or both a heavy or light chain of an immunoglobulin and may, in addition, comprise a third phosphodiester-linked nucleotide or polynucleotide sequence of a nature and function described above.

In yet another aspect, the invention provides a method for producing intact biosynthetic antibody binding sites or native Fv free of all or substantially all constant region amino acids. The method involves enzymatic digestion of chimeric immunoglobulin or at least Fab regions which have been engineered to contain preferential proteolytic cleavage sites located between the variable and constant regions of the immunoglobulin heavy and light chains. Digestion of the intact immunoglobulin with

the appropriate protease yields a complete antigen binding site or Fv fragment. This approach works well in myeloma or hybridoma expression systems.

Accordingly, it is an object of this invention to provide novel proteins comprising biosynthetic antibody binding sites including an amino acid sequence homologous to 5 specific portions of the variable region of immunoglobulin light chain and/or heavy chain, to provide DNA sequences which encode the biosynthetic antibody binding sites, and to provide replicable expression vectors capable of expressing DNA sequences encoding the biosynthetic antibody binding sites. Another object is to provide a generalized method for producing biosynthetic antibody binding site polypeptides of any desired specificity.

10 In its broadest aspect, the invention features a formulation for targeting an epitope on an antigen expressed in a mammal, where the formulation contains a pharmaceutically acceptable carrier in combination with a biosynthetic construct for binding at least one preselected antigen. The dimeric construct has particular utility in diagnostic and therapeutic applications *in vivo*.

15 The invention features the synthesis and use of monomers and dimers of polypeptide constructs belonging to the class of proteins known as single-chain Fv (sFv) polypeptides. The sFv proteins described herein have superior *in vivo* pharmacokinetic properties, including accelerated tissue biodistribution and clearance rates relative to either intact IgG, (Fab)<sub>2</sub> dimers or Fab.

20 The dimeric biosynthetic construct of the invention contains two sFv polypeptide chains defined herein as follows. Each sFv polypeptide chain comprises an amino acid sequence defining at least two polypeptide domains. These domains are connected by a polypeptide linker spanning the distance between the C-terminus of one domain and the N-terminus of the other. The amino acid sequence of each domain includes 25 complementarity determining regions (CDRs) interposed between framework regions (FRs) where the CDRs and FRs of each polypeptide chain together define a binding site immunologically reactive with a preselected antigen. Additionally, each biosynthetic binding site polypeptide chain can have an amino acid sequence peptide bonded and thus contiguous with the C-terminus of each polypeptide chain, referred to herein as a "C-terminal tail" sequence. The term "sFv" refers hereinafter, to an sFv molecule containing 30 such a C-terminal tail sequence. This tail sequence preferably does not contain an  $\alpha$ -helical motif that self-associates with another polypeptide chain of similar sequence but still contains a means for covalently crosslinking two such polypeptide chains together.

When the two sFv' polypeptide chains are crosslinked together, the resulting dimeric construct has a conformation that permits the independent binding of a preselected antigen or antigens to the binding site of each polypeptide chain *in vitro* and *in vivo*. The resulting dimeric constructs have superior *in vivo* pharmacokinetic properties that include significantly enhanced avidity, including enhanced target tissue retention and/or antigen localization properties, as compared with intact IgG, Fab, (Fab)<sub>2</sub> dimers or monomeric sFv.

As will be appreciated by those having ordinary skill in the art, the sequence referred to herein generally as a "C-terminal tail" sequence, peptide bonded to the C-terminus of an sFv and comprising means for crosslinking two sFv polypeptide chains, alternatively may occur at the N-terminus of an sFv ("N-terminal tail") or may comprise part of the polypeptide linker spanning the domains of an individual sFv. The dimeric species created by the crosslinking of sFvs having these alternative "tail" sequences also are contemplated to have a conformation permitting the *in vivo* binding of a preselected antigen by the binding sites of each of the sFv polypeptide chains. Accordingly, descriptions of how to make and use sFv' monomers and dimers comprising a C-terminal tail sequence are extended hereby to include sFv monomers and dimers wherein the tail sequence having crosslinking means occurs at the N-terminus of an sFv or comprises part of the polypeptide linker sequence.

In one embodiment, both polypeptide chains bind the same epitope on a preselected antigen, and the resulting dimeric construct is termed a "homodimer." In another embodiment, the polypeptide chains bind different epitopes on a preselected antigen and the resulting dimeric construct is termed a "heterodimer." In still another embodiment, the two polypeptide chains bind different epitopes on two different, preselected antigens.

The term "epitope", as used herein, refers to a portion of an antigen that makes contact with a particular antibody or antibody analogue. In a typical protein, it is likely that any residue accessible from the surface can form part of one or more antigenic determinants. The term "antigen", as used herein, refers to a molecule that can elicit an immune response and that can react specifically with corresponding antibodies or antibody analogues.

The term "domain", as used herein, refers to an amino acid sequence that folds into a single globular region in its native conformation, and which may exhibit discrete binding or functional properties. The term "polypeptide linker", as used herein, refers to an amino

acid sequence that links the C-terminus of one domain to the N-terminus of the other domain, while still permitting the two domains to maintain their proper physiologically active binding conformations. In a particular aspect of the invention, the currently preferred polypeptide linkers that connect the C-terminus of one domain to the N-terminus 5 of the other domain include part or all of amino acid sequence ((Gly)<sub>4</sub> Ser)<sub>3</sub> set forth in the SEQ. ID. NO.: 7, or ((Ser)<sub>4</sub> Gly)<sub>3</sub> as set forth in SEQ. ID. NO.: 8.

The amino acid sequence of each of the polypeptide domains includes complementarity determining regions interposed between framework regions. The term "complementarity determining regions" or "CDRs", as used herein, refer to amino acid 10 sequences which together define the binding affinity and specificity of the natural Fv region of a native immunoglobulin binding site, or a synthetic polypeptide which mimics this function. CDRs are not necessarily wholly homologous to hypervariable regions of natural Fv molecules, and also may include specific amino acids or amino acid sequences which flank the hypervariable region and have heretofore been considered framework not 15 directly determinative of complementarity. The term "framework regions" or "FRs", as used herein, refers to amino acid sequences which are found naturally occurring between CDRs in immunoglobulins. These FR sequences may be derived in whole or part from the same immunoglobulin as the CDRs, or in whole or part from a different immunoglobulin. For example, in order to enhance biocompatibility of an sFv to be administered to a 20 human, the FR sequences can be derived from a human immunoglobulin and so the resulting humanized sFv will be less immunogenic than a murine monoclonal antibody.

The amino acid sequence of each variable domain includes three CDRs interspersed between four FRs. The two polypeptide domains that define an sFv molecule contain CDRs interspersed between FRs which together form a binding site 25 immunologically reactive with a preselected antigen. The term "immunologically reactive", as used herein, refers to the noncovalent interactions of the type that occur between an immunoglobulin molecule and an antigen for which the immunoglobulin is specific. As used herein, the term "avidity" describes the stability of a complex formed by a multivalent antibody or antibody analogue, with its binding conjugate. Also as used herein, the term "apparent avidity" describes the stability of a complex formed by an antibody or an antibody analogue with its binding conjugate as determined by in vivo 30 immunolocalization studies.

In a preferred aspect of the invention, the CDRs of the polypeptide chain can have an amino acid sequence substantially homologous with at least a portion of the amino acid

sequence of CDRs from a variable region of an immunoglobulin molecule from a first species, together with FRs that are substantially homologous with at least a portion of the amino acid sequence of FRs from a variable region of an immunoglobulin molecule from a second species. Preferably, the first species is mouse and the second species is human.

5 The CDR sequences in the sFv' polypeptides are preferably substantially homologous to an immunoglobulin CDR retaining at least 70%, or more preferably 80% or 90%, of the amino acid sequence of the immunoglobulin CDR, and also retains the immunological binding properties of the immunoglobulin.

10 Each sFv' molecule has a C-terminal polypeptide tail that has a non-self- associating structure and contains at least one crosslinking means. Useful crosslinking means include derivatizable amino acid side chains, particularly those selected from the group consisting of cysteine, lysine, arginine, histidine, glutamate, aspartate, and derivatives and modified forms thereof. In a preferred aspect of the invention, cysteine amino acids are incorporated into the C-terminal tail sequences as the crosslinking means.

15 In another aspect of the invention, the crosslinking means includes one or more amino acids that can be posttranslationally modified. For example, the crosslinking means can include one or more glycosylation sites, wherein the incorporated carbohydrate moieties can be crosslinked in vitro. Preferred glycosylation sequences include Asn-Xaa-Thr and Asn-Xaa-Ser, where Xaa can be any amino acid, wherein the carbohydrate is typically N- linked to asparagine or O-linked to serine or threonine.

20

25 Additionally, the tail also may comprise an amino acid sequence that defines a metal ion chelation motif, and which facilitates purification of the sFv' monomers by metal ion affinity chromatography, such as the IMAC<sup>2+</sup> chromatography system. Furthermore, chelation motifs can be used for binding detectable moieties, such as Technetium<sup>99m</sup> (<sup>99m</sup>Tc) for in vivo imaging. Preferred examples of useful C-terminal tail amino acid sequences wherein the crosslinking means is provided by the sulphydryl group of a cysteine, include: Ser-Cys; (Gly)<sub>4</sub>-Cys; and (His)<sub>6</sub>-(Gly)<sub>4</sub>-Cys; set forth in the Sequence Listing as SEQ. ID. NOS.: 9, 10 and 11, respectively. The (Gly)<sub>4</sub>-Cys sequence facilitates the coordination of <sup>99m</sup>Tc by this tail.

30 In the present invention, monomeric sFv' molecules can be coupled together through the crosslinking means in the C-terminal tails to form either homo- or heterodimeric (sFv')<sub>2</sub> species. The term "sFv coupler", as used herein, refers to the chemical bridge that links two sFv' polypeptide chains together to form a dimeric species. In a preferred aspect of the invention, where the crosslinking means is a cysteine residue,

the linkage is by a disulfide bond. Alternatively, sulphydryl-specific homobifunctional crosslinking reagents, such as bismaleimidohexane, or heterobifunctional crosslinking reagents, can be used to join the two sFv' molecules together. sFv' couplers of preselected length also can be designed to limit interaction between the two sFv' polypeptide chains or 5 to optimize binding of two preselected antigens, including, for example, multiple copies of a receptor expressed on a cell surface in a mammal. An example of such a variable length coupler includes the bismaleimidocaproyl amino acid (MCA) synthetic peptide bridge. Although, in a preferred aspect of the invention a GlySer<sub>3</sub>Gly<sub>2</sub>Ser<sub>3</sub>Lys peptide spacer is used, in theory, any amino acid sequence can be introduced into this type of chemical 10 bridge with a variety of reactive moieties at either end. Consequently, it is possible to design specific linkage groups that can have a predetermined length and flexibility. If a substantially inflexible coupler is desired, then for instance, a polylysine or polyproline peptide may be used. Another benefit of the MCA linkers over many other commercially 15 available linkers is that they are soluble in water. Moreover, the chemical bridge also may be created to enhance the imaging or therapeutic properties of the construct *in vivo* (vide infra). As will be appreciated by those having ordinary skill in the art, the separation distance between, and interaction of, the sFv' monomers in a dimeric construct of the invention also can be modulated by the judicious choice of amino acids in the tail sequences themselves. One of skill in the art can readily test peptide spacers of various 20 lengths and amino acid compositions to select peptide spacers having optimal properties for the particular application.

The dimeric constructs of this invention preferably target a pharmacologically active drug (or other ancillary protein) to a site of interest utilizing the bivalent capability of the dimer. Examples of pharmacologically active drugs include molecules that inhibit 25 cell proliferation and cytotoxic agents that kill cells. The term "cytotoxic agent", as used herein, refers to any molecule that kills cells, and includes anti-cancer therapeutic agents such as doxorubicin. Other, useful molecules include toxins, for instance, the toxic portion of the Pseudomonas exotoxin, phytolaccin, ricin, ricin A chain, or diphtheria toxin, or other related proteins known as ricin A chain-like ribosomal inhibiting proteins, i.e., 30 proteins capable of inhibiting protein synthesis at the level of the ribosome, such as pokeweed antiviral protein, gelonin, and barley ribosomal protein inhibitor.

In such cases, one sFv' can be immunologically reactive with a binding site on an antigen at the site of interest, and the second sFv' in the dimer can be immunologically reactive with a binding site on the drug to be targeted. Alternatively, the construct may

bind one or more antigens at the the site of interest and the drug to be targeted is otherwise associated with the dimer, for example, crosslinked to the chemical bridge itself. The biosynthetic dimeric constructs of this invention also may be used as part of human therapies to target cytotoxic cells such as cytotoxic T-lymphocytes, or pharmacologically active drugs to a preselected site. A bispecific (sFv')<sub>2</sub> heterodimer having specificity for both a tumor antigen and a CD3 antigen, the latter of which is present on cytotoxic T-lymphocytes, thus could mediate antibody dependent cellular cytotoxicity (ADCC) or cytotoxic T-lymphocyte-induced lysis of the tumor cells.

Still another bispecific dimeric construct having cytotoxic properties is a bispecific construct with one sFv' capable of targeting a tumor cell and the second sFv' having catalytic properties that binds an inactive drug, subsequently converting it into an active compound (see for example, U.S. Pat. No. 5,219,732). Such a construct would be capable of inducing the formation of a toxic substance *in situ*. For example, a catalytic sFv' molecule having  $\beta$ -lactamase-like activity can be designed to bind and catalyze the conversion of an inactive lactam derivative of doxorubicin into its active form. Here the bispecific dimer, having binding affinities for both the preselected antigen and the inactive-lactam derivative, is administered to an individual and allowed to accumulate at the desired location. The inactive and nontoxic cytotoxin-lactam derivative then is administered to the individual. Interaction of the derivative with the bispecific (sFv')<sub>2</sub> heterodimer at the site of interest releases the active form of the drug *in situ*, enhancing both the cytotoxicity and specificity of the drug. In this manner the bispecific heterodimers functions to activate a prodrug.

The homo- and heterodimeric biosynthetic constructs also may include a detectable moiety bound either to the polypeptide chain, e.g., to the tail sequence, or to the chemical coupler. The term "detectable moiety", as used herein, refers to the moiety bound to or otherwise complexed with the construct and which can be detected external to, and at a distance from, the site of the complex formation, to permit the imaging of cells or cell debris expressing a preselected antigen. Preferable detectable moieties for imaging include radioactive atoms such as Technetium<sup>99m</sup> (<sup>99m</sup>Tc), a gamma emitter with a half-life of about 6 hours. Non-radioactive moieties useful for *in vivo* magnetic resonance imaging applications include nitroxide spin labels as well as lanthanide and transition metal ions which induce proton relaxation *in situ*. In addition to immunoimaging, the complexed radioactive moieties also may be used in standard radioimmunotherapy protocols to

destroy the targeted cell. Preferable nucleotides for high dose radioimmunotherapy include radioactive atoms such as, <sup>90</sup>Yttrium (<sup>90</sup>Yt), <sup>131</sup>Iodine (<sup>131</sup>I) or <sup>111</sup>Indium (<sup>111</sup>In).

The sFv, sFv' and (sFv')<sub>2</sub> constructs disclosed herein have particular utility as in vivo targeting agents of tumor antigens, including antigens characteristic of breast and ovarian malignancies, such as the c-erbB-2 or c-erbB-2 related antigens. Accordingly, these constructs have particular utility in diagnostic applications as imaging agents of malignant cells, and in therapeutic applications as targeting agents for cytotoxins and other cancer therapeutic agents. In one preferred aspect of the invention, the CDRs of the sFv or sFv' polypeptide chain have an amino acid sequence substantially homologous with the CDRs of the variable region of any one of the following monoclonal antibodies: 741F8, 520C9, and 454C11, all of which bind to c-erbB-2 or c-erbB-2-related antigens. Exemplary sFv' and sFv sequences having CDRs corresponding to the monoclonal antibodies 741F8 and 520C9 are set forth in the Sequence Listing SEQ. ID. NOS.: 1 and 5, respectively.

The term "c-erbB-2" refers to a protein antigen that is an approximately 200 kD acidic glycoprotein having an isoelectric point of about 5.3 and having an extracellular domain overexpressed on the surface of tumor cells, such as breast and ovarian tumor cells in about 25% of cases of breast and ovarian cancer. A "c-erbB-2-related tumor antigen" is a protein located on the surface of tumor cells, such as breast and ovarian tumor cells and which is antigenically related to the c-erbB-2 antigen. That is, the related antigen can be bound by an immunoglobulin that is capable of binding the c-erbB-2 antigen e.g. 741F8, 520C9, and 454C11 antibodies. Related antigens also include antigens comprising an amino acid sequence that is at least 80% homologous, preferably 90% homologous, with the amino acid sequence of c-erbB-2 or an amino acid sequence encoded by a DNA that hybridizes under stringent conditions with a nucleic acid sequence encoding c-erbB-2. As used herein, stringent hybridization conditions are those set forth in Sambrook, et al., 1989, Molecular Cloning; A Laboratory Manual 2nd ed. Cold Spring Harbor Press wherein the hybridization conditions, for example, include 50% formamide, 5x Denhardt's Solution, 5xSSC, 0.1% SDS and 100 µg/ml denatured salmon sperm DNA and the washing conditions include 2xSSC, 0.1% SDS at 37° C. followed by 1xSSC, 0.1% SDS at 68° C. An example of a c-erbB-2-related antigen is the receptor for the epidermal growth factor.

In one embodiment, the biosynthetic antibody binding site is a humanized hybrid molecule which includes CDRs from the mouse 741F8 antibody interposed between FRs derived from one or more human immunoglobulin molecule. The CDRs that bind to the c-erbB-2 epitope can be found in the amino acid residue numbers 31-37, 52-68, 101-110, 5 159-169, 185-191 and 224-233 in SEQ ID NOS.: 1 and 2. The hybrid molecule thus contains binding sites which are highly specific for the c-erbB-2 antigen or c-erbB-2 related antigens held in proper immunochemical binding conformation by human FR amino acid sequences, which are less likely to be recognized as foreign by the human body.

10 The dimeric (sFv')<sub>2</sub> construct can either be homodimeric, wherein the CDR sequences on both monomers define the same binding site, or heterodimeric, wherein the CDR sequences of each sFv' monomer define a different binding site. An example of an (sFv')<sub>2</sub> heterodimer described herein having specificity for both c-erbB-2 and digoxin epitopes can be generated by combining the anti-c-erbB-2 sFv', shown in SEQ. ID. NOS.: 15 1 and 2 with the anti-digoxin sFv', shown in SEQ. ID. NOS.: 3 and 4. The CDRs that bind to the digoxin epitope can be derived from the anti-digoxin murine monoclonal antibody 26-10 (Huston *et al.*, 1988, Proc. Natl. Acad. Sci. USA **85**: 5879-5883) and can be found in the amino acid residue numbers 32 through 36, 48 through 65, 101 through 107, 157 through 170, 188 through 194 and 229 through 234 in the Sequence Listing as SEQ. ID. 20 NOS.: 3 and 4.

Radioimaging or radioimmunotherapy of tumor tissues and malignant cells are preferred aspects of this invention. Overexpression of tumor antigens such as c-erbB-2 and related cell surface antigens in malignant cells allows imaging of the malignant cell or tissue, whether it is well localized, has undergone metastasis or is exposed following cell 25 lysis. The imaging method includes the steps of administering to a mammal a formulation comprising an sFv' or (sFv')<sub>2</sub> dimeric construct having specificity for the antigen tumor and containing a detectable moiety at a concentration sufficient to permit extracorporeal detection of the construct bound to the tumor antigen; and then detecting the biosynthetic construct bound to the tumor antigen. The formulation can be used to particular advantage 30 in gamma scintigraphy or magnetic resonance imaging. Overexpression of c-erbB-2 or related receptors on malignant cells thus allows targeting of sFv' species to the tumor cells, whether the tumor is well-localized or metastatic. In addition, internalization of an sFv-toxin fusion protein permits specific destruction of tumor cells bearing the overexpressed c-erbB-2 or related antigen.

The present invention discloses monomeric and dimeric biosynthetic constructs having enhanced properties as *in vivo* targeting agents when compared with intact monoclonal antibodies or their Fab fragments. The dimeric biosynthetic constructs of the invention also permit the *in vivo* targeting of an epitope on an antigen with greater 5 apparent avidity, including greater tumor specificity, tumor localization and tumor retention properties than that of the Fab fragment having the same CDRs as the construct. Furthermore, the dimeric constructs also permit the *in vivo* targeting of an epitope on an antigen with a greater apparent avidity, including greater tumor localization and tumor retention properties, than either of the monomeric polypeptides individually. Accordingly, 10 the methods and compositions of the present invention provide significant improvements over the prior art with respect to tumor targeting and localization.

The invention also includes methods for producing the homo- and heterodimeric biosynthetic constructs, which include the steps of designing, constructing, expressing, purifying, and refolding the monomeric sFv' polypeptide chains *in vitro*, followed by 15 joining two polypeptide chains together through the crosslinking means in the C-terminal tail sequence, without relying on the tail structure to otherwise assist in dimer formation or enhance transport across a membrane. The invention also includes methods for imaging a preselected antigen in a mammal expressing the preselected antigen. The antigen may be expressed on a cell surface or may be released as part of the cell debris from a dying cell.

20 The foregoing and other objects, features and advantages of the present invention will be made more apparent from the following detailed description of the invention.

#### Brief Description of the Drawings

25 FIG. 1A is a schematic representation of a DNA construct encoding the sFv' biosynthetic binding protein of the invention.

FIG. 1B is a schematic representation of the polypeptide chain encoded by the DNA construct in FIG. 1A.

FIG. 2A is a schematic representation of a refolded sFv' protein in its native conformation.

30 FIG. 2B is a schematic representation showing two folded sFv' polypeptides covalently linked by a disulfide bond.

FIG. 3 is a graphic representation of an *in vitro* competition assay comparing the erbB-2 binding activity of an Fab fragment of the 520C9 monoclonal antibody (filled dots), with that of biosynthetic 520C9 sFv at two different stages of purification: mixture

of folded and unfolded sFv (+) or affinity-purified sFv (squares), and with a material that did not bind to the affinity column (\*).

FIG. 4 lists in tabular form the tumor:organ ratios calculated for various sF and sFv' species injected into tumor-containing mice.

5 FIG. 5 lists in tabular form the percentage of injected dose localized to tumor tissue for various sFv and sFv's species.

10 FIG. 6 is a graphic representation summarizing the comparative tumor retention properties of monomeric and dimeric forms of different sFv' constructs and Fabs represented by bars 1-6. The sFv' species represented by bars 1-5 are based on the V regions of the 741F8 monoclonal antibody. Bar 1 refers to intravenously (i.v.) administered glutathionyl-(sFv'-SerCys) monomer, bar 2 to disulfide linked (sFv'-Gly<sub>4</sub>-Cys)<sub>2</sub>, bar 3 to MCA combined (sFv-Ser-Cys)<sub>2</sub>, bar 4 to BMH cross-linked (sFv-Ser-Cys)<sub>2</sub>, bar 5 to 741F8 Fab and bar 6 to the 26-10 disulfide linked (sFv-Ser-Cys)<sub>2</sub>.

15 FIG. 7 is a schematic representation of an intact IgG antibody molecule containing two light chains, each consisting of one variable and one constant domain, and two heavy chains, each consisting of one variable and three constant domains.

20 FIG. 8A-8D are schematic representations of some of the classes of reagents constructed in accordance with the invention, each of which comprises a biosynthetic antibody binding site.

### Detailed Description of the Invention

It has been discovered that intravenously administered single-chain Fv (sFv) proteins exhibit superior *in vivo* pharmacokinetic properties relative to intact monoclonal antibodies (IgG), (Fab)<sub>2</sub> dimers or Fab fragments. These pharmacokinetic properties include accelerated rates of tissue biodistribution, enhanced target tissue specificity, and exceptionally fast clearance rates. The sFv constructs can be designed to bind to preselected antigens and to have particular utility for *in vivo* immunoimaging and immunotherapy applications. In addition, it also has been discovered that dimeric forms of the constructs, which do not rely on self-associating tail sequences for dimerization or 25 transport across a membrane, can be easily prepared and have improved target tissue localization properties, target tissue retention properties and/or avidity for their targets *in vivo*, relative to monomeric sFv', Fab fragments or intact IgG.

In its broadest aspect, the invention features a formulation for targeting an epitope on an antigen expressed in a mammal. The formulation contains a pharmaceutically

acceptable carrier in combination with a dimeric biosynthetic construct for binding at least one preselected antigen. The preselected antigen either may be an antigen expressed on the surface of a cell or an intracellular component exposed upon lysis of the cell. The sFv, sFv' and (sFv')<sub>2</sub> constructs disclosed herein have particular utility as in vivo targeting agents for detecting malignant cells in a mammal. In a particularly useful embodiment, the constructs disclosed can be used to target the c-erbB-2 or c-erbB-2-related antigens which are overexpressed in certain breast and ovarian cancers. In another embodiment, radioimmunotargeting using radiolabeled (sFv')<sub>2</sub> constructs will be useful for therapeutic as well as diagnostic applications.

10        Provided below are detailed descriptions of biosynthetic sFv, sFv' and (sFv')<sub>2</sub> dimers, useful in the compositions and methods of the invention, together with methods for their construction and administration. Also provided are numerous, non-limiting examples which demonstrate the suitability of these constructs as in vivo targeting reagents for diagnostic and therapeutic applications. More specifically, the examples 15 demonstrate: the construction and expression of sFv polypeptides (Example 1); the renaturation, dimerization and purification of sFv' proteins (Example 2); and the immunoreactivity of the monomeric and dimeric sFv proteins (Example 3).

#### Construction of Biosynthetic Single Chain Fv Proteins.

20        Each of the sFv and sFv' proteins have amino acid sequences that define at least two polypeptide domains. The polypeptide domains are connected by a polypeptide linker spanning the distance between the C-terminus of one domain and the N-terminus of the other. The amino acid sequence of each domain includes complementarity determining regions (CDRs) interposed between framework regions (FRs), where the CDRs and FRs of 25 each polypeptide chain together define a binding site immunologically reactive with a preselected antigen. Preferable polypeptide linkers can be readily tested by one of skill in the art to select a linker of optimal length and amino acid composition.

30        In the case of the sFv' proteins, each polypeptide chain has an additional C-terminal tail amino acid sequence having a substantially non-self-associating structure. More specifically, this is a sequence that does not interact appreciably with a similar sequence under physiological conditions, as is the case for example with the  $\alpha$ -helical leucine zipper motifs found in DNA binding proteins. Each tail sequence also contains a means for crosslinking two such sFv' polypeptide chains together to form an (sFv')<sub>2</sub> dimer.

The resulting (sFv')<sub>2</sub> dimers have conformations which permit the in vivo binding of the preselected antigen by the binding sites of each of the polypeptide chains.

The sFv' constructs of this invention can be further understood by referring to the accompanying Figures 1 and 2. FIG. 1A is a schematic representation of the DNA construct, and FIG. 1B is a schematic representation of the resulting encoded polypeptide chain. FIG. 2 is a schematic representation of the folded sFv' monomer (FIG. 2A) and the dimeric (sFv')<sub>2</sub> construct (FIG. 2B). A single-chain Fv (sFv') polypeptide, shown in FIGS. 1 and 2A, comprises: a heavy chain variable region (V<sub>H</sub>) 10, and a light chain variable region, (V<sub>L</sub>) 14, wherein the V<sub>H</sub> and V<sub>L</sub> domains are attached by polypeptide linker 12. The binding domains defined by V<sub>L</sub> and V<sub>H</sub> include the CDRs 2, 4, 6 and 2', 4', 6', respectively, and FRs 32, 34, 36, 38 and 32', 34', 36', 38', respectively which, as shown in FIG. 2, together define an immunologically reactive binding site or antigenic determinant, 8. Furthermore, the CDRs and FRs may be derived from different immunoglobulins (see infra). The sFv' molecules also contain a C-terminal tail amino acid sequence, 16, comprising an amino acid sequence that will not self-associate with a polypeptide chain having a similar amino acid sequence under physiological conditions, and which contains a means, 18, for the site-directed crosslinking of two such tail sequences. In a currently preferred embodiment, represented in FIGS. 1 and 2, the crosslinking means is the sulphydryl group of a cysteine amino acid. In the monomeric form of the sFv' the crosslinking means, 18, may be blocked by a blocking group, 20. For instance, the blocking group may be a glutathionyl moiety when the crosslinking means, 18, is a cysteine amino acid.

As will be appreciated by those having ordinary skill in the art, the sequence referred to herein generally as a "C-terminal tail" sequence, peptide bonded to the C-terminus of an sFv and comprising means for crosslinking two sFv polypeptide chains, alternatively may occur at the N-terminus of an sFv ("N-terminal tail") or may comprise part of the polypeptide linker spanning the domains of an individual sFv. The dimeric species created by the crosslinking of sFvs having these alternative "tail" sequences also are contemplated to have a conformation permitting the in vivo binding of a preselected antigen by the binding sites of each of the sFv polypeptide chains. Accordingly, descriptions of how to make and use sFv' monomers and dimers comprising a C-terminal tail sequence are extended hereby to include sFv monomers and dimers wherein the tail sequence having crosslinking means occurs at the N-terminus of an sFv or comprises part of the polypeptide linker sequence.

The CDR and FR polypeptide segments are designed empirically based on sequence analysis of Fv regions of preexisting antibodies, such as those described in U.S. Pat. No. 4,753,894, hereby incorporated by reference. Numerous examples of sFv polypeptide chains now exist in the art and are summarized in Huston *et al.*, 1993, *Intern. Rev. Immunol.* 10: 195-217, hereby incorporated by reference.

5 The sFv and sFv' polypeptide chains of the invention are biosynthetic in the sense that they are synthesized, transfected into a cellular host, and protein expressed from a nucleic acid containing genetic sequences based in part on synthetic DNA. Synthetic DNA is understood to include recombinant DNA made by ligation of fragments of DNA 10 derived from the genome of a hybridoma, mature B cell clones, a cDNA library derived from natural sources, or by ligation of plural, chemically synthesized oligonucleotides. The proteins of the invention are properly characterized as "antibody binding sites", in that these synthetic single polypeptide chains are able to refold into a 3-dimensional conformation with specificity and affinity for a preselected epitope on an antigen.

15 A detailed description for engineering and producing sFv proteins by recombinant means appears in U.S. Pat. No. 5,091,513 claiming priority from U.S.S.N. 052,800, filed May 21, 1987, assigned to Creative BioMolecules, Inc., hereby incorporated by reference. The polypeptide chains of the invention are antibody-like in that their structure is 20 patterned after regions of native antibodies known to be responsible for antigen recognition.

25 As is now well known, Fv, the minimum antibody fragment which contains a complete antigen recognition and binding site, consists of a dimer of one heavy and one light chain variable domain in tight, noncovalent association. It is in this configuration that the three complementarity determining regions of each variable domain interact to define an antigen binding site on the surface of the  $V_H$  - $V_L$  dimer. Collectively, the six 30 complementarity determining regions confer antigen binding specificity to the antibody. FRs flanking the CDRs have a tertiary structure which is essentially conserved in native immunoglobulins of species as diverse as human and mouse. These FRs serve to hold the CDRs in their appropriate orientation. The constant domains are not required for binding function, but may aid in stabilizing  $V_H$  - $V_L$  interaction. Even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than an entire binding site (Painter *et al.*, 1972, *Biochem.* 11: 1327-1337).

This knowledge of the structure of immunoglobulin proteins has now been exploited to develop biosynthetic antibody binding sites provided by this invention.

The biosynthetic antibody binding sites embodying the invention are biosynthetic in the sense that they are synthesized in a cellular host made to express a synthetic DNA, 5 that is, a recombinant DNA made from ligation of plural, chemically synthesized oligonucleotides, or by ligation of fragments of DNA derived from the genome of a hybridoma, mature B cell clone, or a cDNA library derived from such natural sources. The proteins of the invention are properly characterized as "antibody binding sites" in that these synthetic molecules are designed specifically to have at least some affinity for a 10 preselected antigenic substance. The polypeptides of the invention are antibody-like in that their structure is patterned after regions of native antibodies known to be responsible for antigen recognition.

More specifically, the structure of these biosynthetic proteins in the region which 15 impart the binding properties to the protein, is analogous to the Fv region of a natural antibody. It comprises a series of regions consisting of amino acids defining at least three polypeptide segments which together form the tertiary molecular structure responsible for affinity and binding. These regions are herein called complementarity determining regions or CDRs. These CDR regions are held in appropriate conformation by 20 polypeptide segments analogous to the framework regions of the Fv fragment of natural antibodies.

The term CDR, as used herein, refers to amino acid sequences which together 25 define the binding affinity and specificity of the natural Fv region of a native immunoglobulin binding site, or a synthetic polypeptide which mimics this function. CDRs typically are not wholly homologous to hypervariable regions of natural Fvs, but rather also include specific amino acids or amino acid sequences which flank the hypervariable region and have heretofore been considered framework not directly determinitive of complementarity. The term FR, as used herein, refers to amino acid sequences interposed between CDRs.

The CDR and FR polypeptide segments are designed empirically based on 30 sequence analysis of the Fv region of preexisting antibodies or of the DNA encoding them. In one embodiment, the amino acid sequences constituting the FR regions are analogous to the FR sequences of a first preexisting antibody, for example, a human IgG. The amino acid sequences constituting the CDR regions are analogous to the sequences from a second, different preexisting antibody, for example, the CDRs of a murine IgG.

Alternatively, the CDRs and FRs from a single preexisting antibody from, e.g., an unstable or hard to culture hybridoma, may be copied in their entirety.

Practice of the invention enables the design and biosynthesis of various reagents, all of which are characterized by a region having affinity for a preselected antigenic substance. Other regions of the biosynthetic protein are designed with the particular planned utility of the protein in mind. Thus, if the reagent is designed for intravascular use in mammals, the FR regions comprise amino acids similar or identical to at least a portion of the framework region amino acids of antibodies native to that mammalian species. On the other hand, the amino acids comprising the CDRs may be analogous to a portion of the amino acids from the hypervariable region (and certain flanking amino acids) of an antibody having a known affinity and specificity, e.g., a murine or rat monoclonal antibody.

Other sections, e.g.,  $C_H$  and  $C_L$ , of native immunoglobulin protein structure need not be present and normally are intentionally omitted from the biosynthetic proteins of this invention. However the BABS of the invention may comprise additional polypeptide regions defining a bioactive region, e.g., a toxin or enzyme, or a site onto which a toxin or a remotely detectable substance can be attached.

The clinical administration of the BABS of the invention, which display the activity of native, relatively small  $F_v$ ,  $V_H$ , or  $V_L$  fragments, affords a number of advantages over the use of larger fragments or entire antibody molecules. The BABS of this invention offer fewer cleavage sites to circulating proteolytic enzymes and thus offer greater stability. They reach their target tissue more rapidly, and are cleared more quickly from the body. They also have reduced immunogenicity. In addition, their smaller size facilitates coupling to other molecules in drug targeting and imaging application.

The invention thus provides intact biosynthetic antibody binding sites analogous to  $V_H$ - $V_L$  dimers, either non-covalently associated, disulfide bonded, or linked by a polypeptide sequence to form a composite  $V_H$ - $V_L$  or  $V_L$ - $V_H$  polypeptide which is essentially free of the remainder of the antibody molecule. The invention also provides proteins analogous to an independent  $V_H$  or  $V_L$  domain. Any of these proteins may be provided in a form linked to amino acid sequences exclusive of those of the variable domain, for example, to amino acids analogous or homologous to proteins of a constant domain, or another bioactive molecules such as a hormone or toxin. A proteolytic cleavage site can also be designed into the region separating the variable region-like

sequences from other pendant sequences so as to facilitate cleavage of intact  $V_H$  and/or  $V_L$ , free of other protein.

FIGS. 8A, 8B, 8C, and 8D illustrate four examples of protein structures embodying the invention that can be produced by following the teaching disclosed herein. All are characterized by one or two biosynthetic polypeptide segments defining a binding site 3, and comprising amino acid sequences comprising CDRs and FRs, often derived from different immunoglobulins, or sequences homologous to a portion of CDRs and FRs from different immunoglobulins. FIG. 8A depicts a single chain Fv comprising a polypeptide 10 having an amino acid sequence analogous to the variable region of an immunoglobulin heavy chain, bound through its carboxyl end to a polypeptide linker 12, which in turn is bound to a polypeptide 14 having an amino acid sequence analogous to the variable region of an immunoglobulin light chain. Of course, the light and heavy chain domains may be in reverse order. The linker 12 should be at least long enough (e.g., about 15 amino acids or about 40A) to permit the chains 10 and 14 to assume their proper conformation. The linker 12 may comprise an amino acid sequence homologous to a sequence identified as "self" by the species into which it will be introduced, if drug use is intended. Unstructured, hydrophilic amino acid sequences are preferred. It may also comprise a bioactive polypeptide such as a cell toxin which is to be targeted by the binding site, or a segment easily labeled by a radioactive reagent which is to be delivered, e.g., to the site of a tumor comprising an epitope recognized by the binding site. Other proteins or polypeptides may be attached to either the amino or carboxyl terminus of protein of the type illustrated in FIG. 8A. As an example, a helically coiled polypeptide structure illustrating a leader comprising a protein A fragment is shown extending from the amino terminal end of  $V_H$  domain 10.

FIG. 8B illustrates two separate chains non-covalently associated and defining a binding site 3. It comprises separate peptides 16 and 18 comprising a chimeric  $V_H$  and  $V_L$  of the type described above. The carboxyl terminus of each protein chain may be designed to include one or more cysteine residues so that oxidation of properly folded structures produces disulfide bonds (see FIG. 8C) further stabilizing the BABS. Either or both of the polypeptides may further comprise a fused protein imparting other biological properties to the reagent in addition to the ability to bind to the antigen as specified by the interaction of the triplet CDRs on the respective polypeptides 16 and 18.

FIG. 8D depicts another type of reagent, comprising only one set of three CDRs, e.g., analogous to a heavy chain variable region, which retains a measure of affinity for the

antigen. Attached to the carboxyl end of the polypeptide comprising the FR and CDR sequences constituting the binding site 3 is a Pendant Protein P consisting of, for example, a toxin, therapeutic drug, binding protein, enzyme or enzyme fragment, site of attachment for an imaging agent (e.g., to chelate a radioactive ion such as Indium), or site of attachment to an immobilization matrix so that the BABS can be used in affinity chromatography.

Of course, the protein may comprise more than one binding site or copies of a single binding site, and a number of other functional regions.

As is evidenced from the foregoing, the invention provides a large family of reagents comprising proteins, at least a portion of which defines a binding site patterned after the variable region or regions of natural immunoglobulins. It will be apparent that the nature of any protein fragments linked to the BABS, and used for reagents embodying the invention, are essentially unlimited, the essence of the invention being the provision, either alone or linked in various ways to other proteins, of binding sites having specificities to any antigen desired.

The BABS of the invention are designed at the DNA level. The chimeric or synthetic DNAs are then expressed in a suitable host system, and the expressed proteins are collected and renatured if necessary.

The ability to design the BABS of the invention depends on the ability to determine the sequence of the amino acids in the variable region of monoclonal antibodies of interest, or the DNA encoding them. Hybridoma technology enables production of cell lines secreting antibody to essentially any desired substance that produces an immune response. RNA encoding the light and heavy chains of the immunoglobulin can then be obtained from the cytoplasm of the hybridoma, and the 5' end portion of the mRNA can be used to prepare the cDNA for subsequent sequencing, or the amino acid sequence of the hypervariable and flanking framework regions can be determined by amino acid sequencing of the H and L chains and their V region fragments. Such sequence analysis is now conducted routinely. This knowledge permits one to design synthetic genes encoding FR and CDR sequences which likely will bind the antigen. These synthetic genes are then prepared using known techniques, or using the technique disclosed below, and then inserted into a suitable host, expressed, and purified. Depending on the host cell, renaturation techniques may be required to attain proper conformation. The various proteins are then tested for binding ability, and one having appropriate affinity is selected for incorporation into a reagent of the type described above. If necessary, point

substitutions seeking to optimize binding may be made in the DNA using conventional cassette mutagenesis or other protein engineering methodology.

Of course, the processes for manipulating, amplifying, and recombining DNA which encode amino acid sequences of interest are generally well known in the art, and therefore, not described in detail herein. Methods of identifying and isolating genes encoding antibodies of interest are well understood, and described in the patent and other literature. In general, the methods involve selecting genetic material coding for amino acids which define the CDRs and FRs of interest according to the genetic code.

Accordingly, the construction of DNAs encoding BABS as disclosed herein can be done using known techniques involving the use of various restriction enzymes which make sequence specific cuts in DNA to produce blunt ends or cohesive ends, DNA ligases, techniques enabling enzymatic addition of sticky ends to blunt-ended DNA, construction of synthetic DNAs by assembly of short or medium length oligonucleotides, cDNA synthesis techniques, and synthetic probes for isolating immunoglobulin genes. Various promoter sequences and other regulatory DNA sequences used in achieving expression, and various types of host cells are also known and available. Conventional transfection techniques, and equally conventional techniques for cloning and subcloning DNA are useful in the practice of this invention and known to those skilled in the art. Various types of vectors may be used such as plasmids and viruses including animal viruses and bacteriophages. The vectors may exploit various marker genes which impart to a successfully transfected cell a detectable phenotypic property that can be used to identify which of a family of clones has successfully incorporated the recombinant DNA of the vector.

One method for obtaining DNA encoding the BABS disclosed herein is by assembly of synthetic oligonucleotides produced in a conventional, automated, polynucleotide synthesizer followed by ligation with appropriate ligases. For example, overlapping, complementary DNA fragments comprising 15 bases may be synthesized semi manually using phosphoramidite chemistry, with end segments left unphosphorylated to prevent polymerization during ligation. One end of the synthetic DNA is left with a "sticky end" corresponding to the site of action of a particular restriction endonuclease, and the other end is left with an end corresponding to the site of action of another restriction endonuclease. Alternatively, this approach can be fully automated. The DNA encoding the BABS may be created by synthesizing longer single strand fragments (e.g.,

50-100 nucleotides long) in, for example, a Biosearch oligonucleotide synthesizer, and then ligating the fragments.

Still another method of producing the BABS of the invention is to produce a synthetic DNA encoding a polypeptide comprising, e.g., human FRs, and intervening "dummy" CDRs, or amino acids having no function except to define suitably situated unique restriction sites. This synthetic DNA is then altered by DNA replacement, in which restriction and ligation is employed to insert synthetic oligonucleotides encoding CDRs defining a desired binding specificity in the proper location between the FRs.

This technique is dependent upon the ability to cleave a DNA corresponding in structure to a variable domain gene at specific sites flanking nucleotide sequences encoding CDRs. These restriction sites in some cases may be found in the native gene. Alternatively, non-native restriction sites may be engineered into the nucleotide sequence resulting in a synthetic gene with a different sequence of nucleotides than the native gene, but encoding the same variable region amino acids because of the degeneracy of the genetic code. The fragments resulting from endonuclease digestion, and comprising FR-encoding sequences, are then ligated to non-native CDR-encoding sequences to produce a synthetic variable domain gene with altered antigen binding specificity. Additional nucleotide sequences encoding, for example, constant region amino acids or a bioactive molecule may also be linked to the gene sequences to produce a bifunctional protein.

The expression of these synthetic DNAs can be achieved in both prokaryotic and eucaryotic systems via transfection with the appropriate vector. In *E. coli* and other microbial hosts, the synthetic genes can be expressed as fusion protein. Expression in eucaryotes can be accomplished by the transfection of DNA sequences encoding CDR and FR region amino acids into a myeloma or other type of cell line. By this strategy intact hybrid antibody molecules having hybrid Fv regions and various bioactive proteins including a biosynthetic binding domain may be produced. For fusion protein expressed in bacteria subsequent proteolytic cleavage of the isolated  $V_H$  and  $V_L$  fusions can be performed to yield free  $V_H$  and  $V_L$ , which can be renatured, and reassociated (or used separately) to obtain an intact biosynthetic, hybrid antibody binding site.

Heretofore, it has not been possible to cleave the heavy and light chain region to separate the variable and constant regions of an immunoglobulin so as to produce intact Fv, except in specific cases not of general utility. However, one method of producing BABS in accordance with this invention is to redesign an immunoglobulin at the DNA level so as to alter its specificity and so as to incorporate a cleavage site and "hinge

region" between the variable and constant regions of both the heavy and light chains. Such chimeric antibodies can be produced in transfectomas or the like and subsequently cleaved using a preselected endopeptidase. The engineering principles involved in these easily cleaved constructs are disclosed in detail in copending U.S. application Ser. No. 5 028,484 filed Mar. 20, 1987 by Huston et al.

The hinge region is a sequence of amino acids which serve to promote efficient cleavage by a preselected cleavage agent at a preselected, built-in cleavage site. It is designed to promote cleavage preferentially at the cleavage site when the polypeptide is treated with the cleavage agent in an appropriate environment.

10 The hinge can take many different forms. Its design involves selection of amino acid residues (and a DNA fragment encoding them) which impart to the region of the fused protein about the cleavage site an appropriate polarity, charge distribution, and stereochemistry which, in the aqueous environment where the cleavage takes place, efficiently exposes the cleavage site to the cleavage agent in preference to other potential 15 cleavage sites that may be present in the polypeptide, and/or to improve the kinetics of the cleavage reaction. In specific cases the amino acids of the hinge are selected and assembled in sequence based on their known properties, and then the fused polypeptide sequence is expressed, tested, and altered for empirical refinement.

20 The hinge region is free of cysteine. This enables the cleavage reaction to be conducted under conditions in which the protein assumes its tertiary conformation, and 25 may be held in this conformation by intramolecular disulfide bonds. It has been discovered that in these conditions access of the protease to potential cleavage sites which may be present within the target protein is hindered. The hinge region may comprise an amino acid sequence which includes one or more proline residues. This allows formation of a substantially unfolded molecular segment. Aspartic acid, glutamic acid, arginine, lysine, serine, and threonine residues maximize ionic interactions and may be present in amounts and/or in sequence which renders the moiety comprising the hinge water soluble.

30 In the case of single chain Fv comprising fused H and L chains, the cleavage site preferably is immediately adjacent the Fv polypeptide and comprises one or a sequence of amino acids exclusive of any one or sequence found in the amino acid structure of the BABS. Where BABS  $V_H$  and  $V_L$  regions are on separate chains (i.e., see FIG. 7), the cleavage sites may be either immediately adjacent their C-terminal ends, thereby releasing Fv dimer of  $V_H$  and  $V_L$  upon appropriate cleavage (i.e., to yield the species of FIG. 8B), or

may follow pendant polypeptides with or without cysteine that yield, respectively, the species of FIG. 8C or 8D upon digestion.

The cleavage site preferably is designed for cleavage by a specific selected agent. Endopeptidases are preferred, although non-enzymatic (chemical) cleavage agents may be used. Many useful cleavage agents, for instance, cyanogen bromide, dilute acid, trypsin, *Staphylococcus aureus* V-8 protease, post proline cleaving enzyme, blood coagulation Factor Xa, enterokinase, and renin, recognize and preferentially or exclusively cleave particular cleavage sites. One currently preferred cleavage agent is V-8 protease. The currently preferred cleavage site is a Glu residue. Other useful enzymes recognize multiple residues as a cleavage site, e.g., factor Xa (Ile-Glu-Gly-Arg) or enterokinase (Asp-Asp-Asp-Asp-Lys).

With the help of a computer program and known variable region DNA sequences, synthetic  $V_L$  and  $V_H$  genes may be designed which encode native or near native FR and CDR amino acid sequences from an antibody molecule, each separated by unique restriction sites located as close to FR-CDR and CDR-FR borders as possible. Alternatively, genes may be designed which encode native FR sequences which are similar or identical to the FRs of an antibody molecule from a selected species, each separated by "dummy" CDR sequences containing strategically located restriction sites. These DNAs serve as starting materials for producing BABS, as the native or "dummy" CDR sequences may be excised and replaced with sequences encoding the CDR amino acids defining a selected binding site. Alternatively, one may design and directly synthesize native or near-native FR sequences from a first antibody molecule, and CDR sequences from a second antibody molecule. Any one of the  $V_H$  and  $V_L$  sequences described above may be linked together directly, either via an amino acids chain or linker connecting the C-terminus of one chain with the N-terminus of the other, or via C-terminal cysteine residues on each of the  $V_H$  and  $V_L$ .

These genes, once synthesized, may be cloned with or without additional DNA sequences coding for, e.g., an antibody constant region, or a leader peptide which facilitates secretion or intracellular stability of a fusion polypeptide. The genes then can be expressed directly in an appropriate host cell, or can be further engineered before expression by the exchange of FR, CDR, or "dummy" CDR sequences with new sequences. This manipulation is facilitated by the presence of the restriction sites which have been engineered into the gene at the FR-CDR and CDR-FR borders.

The engineered genes can be expressed in appropriate prokaryotic hosts such as various strains of *E. coli*, and in eucaryotic hosts such as Chinese hamster ovary cell, mouse myeloma, and human myeloma/transfected cells.

For example, if the gene is to be expressed in *E. coli*, it may first be cloned into an expression vector. This is accomplished by positioning the engineered gene downstream from a promoter sequence such as Trp or Tac, and a gene coding for a leader peptide such as fragment B of protein A (FB). The resulting expressed fusion protein accumulates in refractile bodies in the cytoplasm of the cells, and may be harvested after disruption of the cells by French press or sonication. The refractile bodies are solubilized, and the expressed proteins refolded and cleaved by the methods already established for many other recombinant proteins.

If the engineered gene is to be expressed in myeloma cells, the conventional expression system for immunoglobulins, it is first inserted into an expression vector containing, for example, the Ig promoter, a secretion signal, immunoglobulin enhancers, and various introns. This plasmid may also contain sequences encoding all or part of a constant region, enabling an entire part of a heavy or light chain to be expressed. The gene is transfected into myeloma cells via established electroporation or protoplast fusion methods. Cells so transfected can express  $V_L$  or  $V_H$  fragments,  $V_L - V_H$  heterodimers,  $V_H - V_L$  or  $V_L - V_H$  single chain polypeptides, complete heavy or light immunoglobulin chains, or portions thereof, each of which may be attached in the various ways discussed above to a protein domain having another function (e.g., cytotoxicity).

Vectors containing a heavy chain V region (or V and C regions) can be cotransfected with analogous vectors carrying a light chain V region (or V and C regions), allowing for the expression of noncovalently associated Fvs (or complete antibody molecules).

The single-chair polypeptide chains of the invention are first derived at the DNA level. The sFv DNAs are preferably expressed in *E. coli*, the resulting polypeptide chains being solubilized from inclusion bodies, refolded *in vitro*, labeled with a detectable moiety, such as  $^{99m}\text{Tc}$ , and dimerized to form a biosynthetic  $(\text{sFv}')_2$  construct. Of course, the constructs disclosed herein may also be engineered for secretion from the host cell, for example, secretion into the periplasmic space of an *E. coli* cell, as described by Pack and Pluckthun, (*Biochem.*, 1992, 31: 1579-1584), or into the culture supernatant of a mammalian cell (for example, as described by Traunecker, *et al.*, 1991, *EMBO J.* 10: 3655-3659).

The ability to design the single polypeptide chains of the invention depends on the ability to identify Fv binding domains of interest, and to obtain the DNA encoding these variable regions. Hybridoma technology enables the production of cell lines that secrete antibodies to essentially any desired substance that elicits an immune response. For 5 example, U.S. Pat. No. 4,753,894 describes some monoclonal antibodies of interest which recognize c-erbB-2 related antigens on breast cancer cells, and explains how such antibodies were obtained. One monoclonal antibody that is particularly useful in targeting the c-erbB-2 antigen is 741F8 (Bjorn *et al.*, 1985, Cancer Res. 45: 1214-1221; U.S. Pat. 10 4,753,894). This antibody specifically recognizes the c-erbB-2 antigen expressed on the surface of various tumor cell lines, and exhibits very little binding to normal tissues. Other monoclonal antibodies that bind c-erbB-2 or related antigens include 520C9 and 454C11 (Frankel *et al.*, 1985, J. Biol. Resp. Modif. 4: 273-286; Ring *et al.*, 1989, Cancer Res. 49: 3070-3080, Ring *et al.*, 1991, Molec. Immunol. 28: 915-917; U.S. Pat. No. 4,753,894 and 15 5,169,774). sFv' sequences with the desired specificity can also be derived from phage antibody cloning of combinatorial V gene libraries. Such sequences could be based on cDNA derived from mice preimmunized with tumor cell membranes bearing c-erbB-2 or related antigenic fragments, (See, for example, Clackson *et al.*, (1991) Nature 352: 624-628).

The process of designing DNA encoding the single polypeptide chain of interest 20 can be accomplished as follows. Either synthetic DNA duplexes can be ligated together to form a synthetic gene or relevant DNA fragments can be cloned from libraries. In the latter procedure, mRNA encoding the light and heavy chains of the desired immunoglobulin may be isolated from hybridomas producing the immunoglobulin and reverse transcribed into cDNA. The V<sub>H</sub> and V<sub>L</sub> genes subsequently can be isolated by 25 standard procedures, for instance, by colony hybridization of cDNA libraries (see for example, Sambrook *et al.*, eds., 1989, Molecular Cloning, Cold Spring Harbor Laboratories Press, N.Y.) or by polymerase chain reaction (PCR) (see for example, Innis *et al.*, eds., 1990, PCR Protocols, A guide to methods and applications, Academic Press). Both procedures are well known in the art.

30 Still another approach involves the design and construction of synthetic variable domain genes encoding a predetermined, specific Fv binding site. For example, with the help of a computer program, such as Compugene, one may design and directly synthesize native or near-native FR sequences from a first antibody molecule, and CDR sequences from a second antibody molecule. The resulting V<sub>H</sub> and V<sub>L</sub> gene sequences can then be

genetically linked together by means of a linker connecting the C-terminus of one chain with the N-terminus of the other.

Practice of the invention enables the design and synthesis of various single-chain binding proteins, all of which are characterized by a region having affinity for a preselected epitope on an antigen. Other regions of the biosynthetic protein are designed with the particular planned utility of the protein in mind. Thus, if the reagent is designed for intravascular use in mammals, the FRs may include amino acid sequences which are similar or identical to at least a portion of the FR amino acid sequences of antibodies native to that species. The amino acid sequences constituting the CDRs may be analogous to the sequences from a second, different preexisting antibody having specificity for the antigen of interest (e.g. a murine or other human IgG). Alternatively, the CDRs and FRs may be copied in their entirety from a single pre-existing monoclonal antibody cell line or a desirable sFv species may be cloned from a repertoire library derived from preimmunized or naive animals.

It is noted however, that the linear arrangement of the  $V_L$  and  $V_H$  domains in the DNA sequence of Fig. 1 is not critical. That is, although the sequence represented in Fig. 1A encodes a heavy chain variable region followed by the light chain variable region, as will be appreciated by those skilled in the art, the sFv may be constructed so that the light and heavy chain domains are in reverse order.

As mentioned above, the  $V_H$  and  $V_L$  domains of the sFv are linked in the gene construct by means of a linker 12 (FIG. 1A). The linker should be at least long enough (e.g., about 10 to 15 amino acids or at least 40 Angstroms in length) to permit domains 10 and 14 to assume their proper conformations and interdomain relationships. The linkers preferably comprise hydrophilic amino acids that assume an unstructured configuration under physiological conditions, and are free of residues having large side groups that could interfere with proper folding of the  $V_H$ ,  $V_L$ , or pendant chains. Examples of currently preferred linkers include either part or all of the amino acid sequences  $((Gly)_4Ser)_3$  and  $((Ser)_4Gly)_3$ , set forth in the Sequence Listing as SEQ. ID. NOS.: 7 and 8, respectively. The linker may also include an amino acid sequence homologous to a sequence identified as "self" by the species into which it will be introduced, particularly if a therapeutic application is intended.

Considerations for Suitable C-terminal Tail Sequences.

As mentioned above, the sFv' polypeptide chains further comprise a C-terminal tail containing at least one amino acid that can be derivatized or posttranslationally modified to enable crosslinking of two such sFv' monomers. In preferred aspects of the invention, the tail sequences include one or more of the sequences Ser-Cys, (Gly)<sub>4</sub>-Cys and (His)<sub>6</sub>-  
5 (Gly)<sub>4</sub>-Cys, set forth in the Sequence Listing as SEQ. ID. NOS.: 9, 10, and 11, respectively. The C-terminal tails preferably do not form  $\alpha$ -helical structures which self-  
associate under physiological conditions, such as the  $\alpha$ -helical leucine zipper motifs found  
in DNA binding proteins (O'Shea *et al.*, 1989, Science **243**: 538-542, O'Shea *et al.*, 1991,  
10 Science **254**: 539-544) or the four-helix bundle motifs found in recombinant ion channels  
(Hill *et al.*, 1990, Science **294**: 543-546).

Suitable derivatizable amino acid side chains may be selected from the group consisting of cysteine, lysine, arginine, histidine, glutamate, aspartate and derivatives or modified forms thereof. In a preferred aspect of the invention, cysteine amino acids are incorporated into the C-terminal tail sequences as the crosslinking means.

15 Also envisioned to be useful are posttranslationally modified amino acids that can be crosslinked *in vitro*. More specifically, the glycosyl moieties present on glycosylated amino acids, following secretion out of the cell, can be covalently attached *in vitro* using bifunctional linkers on standard sugar chemistry (see for example, E. A. Davidson (1967) Carbohydrate Chemistry, Holt, Kinehart and Winston, N.Y.; W. J. Lennarz (1980) The  
20 Biochemistry of Glycoproteins and Proteoglycans, Plenum Press, N.Y.). Particularly useful glycosylation sites include the sequences Asn-Xaa-Thr and Asn-Xaa-Ser, wherein Xaa is any amino acid. Where crosslinking of glycosyl moieties is contemplated, the glycosylation sequences need not include a cysteine.

25 The tail also may comprise an amino acid sequence defining an ion chelation motif which can be used as part of a purification protocol for isolating of the sFv' monomers by metal ion affinity chromatography (e.g., by means of a (His)<sub>6</sub> tail on an IMAC chromatography column), as well as for chelating ions of detectable moieties such as Technetium<sup>99m</sup> or <sup>111</sup>Indium for *in vivo* imaging applications.

30 sFv' Coupler Considerations.

In the present invention, two monomeric sFv' proteins are crosslinked together through their C-terminal tails to form an (sFv')<sub>2</sub> dimer. The term "sFv coupler", as used

herein, refers to chemical bridges that join the crosslinking residues in each of the sFv' molecules.

In one preferred aspect of the invention, where the crosslinking residue is a cysteine residue, the chemical bridge can be a disulfide bond. Alternatively, sulfhydryl-specific crosslinking reagents can be used to join two sFv' molecules together. An example of such a cysteine-specific chemical bridge includes the bifunctional crosslinking reagent bismaleimidohexane (BMH), a water insoluble linker that can be obtained from Pierce, Rockford, Ill. Other bifunctional crosslinking agents include heterobifunctional crosslinkers which can be used to join two sFv' species together where the crosslinking residues in each of the sFv' C-terminal tail sequences are different, such as, a C-terminal cysteine on one sFv' and a C-terminal lysine on the other. Useful heterobifunctional crosslinking agents include 4-succinimidylloxycarbonyl-methyl-(2-pyridylidithio)-toluene (SMPT) or N-succinimidyl-3-(2-pyridylidithio)propionate (SPDP), both of which can be obtained from Pierce, Rockland, Ill.

sFv' couplers of variable length also can be prepared to limit steric interaction of two coupled sFv' proteins. An example of such an sFv' coupler includes a peptide bridge, such as the water soluble bismaleimidocaproyl amino acid (MCA) linker. Although in a preferred aspect of the invention, an MCA-GlySer<sub>3</sub>Gly<sub>2</sub>Ser<sub>3</sub>Lys-MCA linker is used, in theory, any amino acid sequence can be introduced into this type of chemical bridge-spacer group.

Suitable MCA-peptide chemical bridges can be synthesized on polystyrene resins functionalized with hydroxymethylphenoxyacetic acid (HMP) to allow formation of free acids at the C-terminus following deblocking. During the synthesis of the preferred peptide sequence Gly-Ser<sub>3</sub>-Gly<sub>2</sub>-Ser<sub>3</sub>-Lys the C-terminal lysine is esterified to the resin and other amino acids are added as N- $\alpha$ -Fmoc protected derivatives. DIC/hydroxybenzotriazol activated amino acids are coupled for 90 minutes after which the N- $\alpha$ -Fmoc protected groups are deprotected with 20% piperidine in dimethylformamide (DMF). Upon completion of the synthesis, the peptide is cleaved from the resin and deprotected with 95% trifluoroacetic acid (TFA) in water. The crude peptide then is dissolved in 0.1M phosphate buffer pH 7 and reacted overnight at room temperature with maleimidocaproic acid N-hydroxysuccinimide ester. The resulting homobifunctional peptide crosslinker can be purified by reverse-phase HPLC, for example, on a Vydac 1 x 25 cm column using acetonitrile/water/TFA buffers.

With this procedure, it is possible to generate linkers having specific lengths and flexibilities. Since polypeptides having particular secondary structures and flexibilities are well documented in the art, it is possible to judiciously design the peptide couplers with optimal length and flexibility to optimize binding to two preselected antigens on a cell surface. As will be appreciated by those skilled in the art, the separation distance between, and interaction of, the sFv' monomers in a dimeric construct of the invention also can be modulated by the judicious choice of amino acids in the tail sequences themselves.

#### Dimer Considerations.

Using the approaches described above,  $(sFv')_2$  dimers readily can be prepared wherein the resulting dimers either can be homodimeric, where the CDR sequences define the same epitope binding site, or heterodimeric, where the CDR sequences of each sFv' monomer define different epitope binding sites.

The dimeric constructs of this invention preferably target a pharmacologically active drug (or other ancillary protein) to a site of interest utilizing the bivalent capability of the dimer. Examples of pharmacologically active drugs include molecules that inhibit cell proliferation and cytotoxic agents that kill cells. Other, useful molecules include toxins, for instance, the toxic portion of the *Pseudomonas* exotoxin, phytolaccin, ricin, ricin A chain, or diphtheria toxin, or other related proteins known as ricin A chain-like ribosomal inhibiting proteins, i.e., proteins capable of inhibiting protein synthesis at the level of the ribosome, such as pokeweed antiviral protein, gelonin, and barley ribosomal protein inhibitor.

In such cases, one sFv' can be immunologically reactive with a binding site on an antigen at the site of interest, and the second sFv' in the dimer can be immunologically reactive with a binding site on the drug to be targeted. For example, the  $(sFv')_2$  dimers may have specificity for both c-erbB-2 and a pharmacologically active drug or cytotoxic agent. The resulting dimer can thus target the agent or drug to tissues expressing the c-erbB-2 antigen in vivo. Alternatively, the construct may bind one or more antigens at the site of interest and the drug to be targeted is otherwise associated with the dimer, for example, by crosslinking to the chemical bridge itself.

Other bispecific  $(sFv')_2$  constructs having particular utility in targeting malignant cells, include constructs wherein one has specificity for a c-erbB-2 or related tumor antigen, and the second determinant has specificity for a different cell surface protein, such as the CD3 antigen found on cytotoxic T-lymphocytes. The heterodimeric  $(sFv')_2$

construct then could mediate antibody dependent cellular cytotoxicity (ADCC) or cytotoxic T-lymphocyte-induced lysis of the tumor cells expressing the c-erbB-2 antigen.

Still another bispecific dimeric construct having cytotoxic properties is a bispecific construct with one sFv' capable of targeting a tumor cell and the second being a catalytic sFv' that binds an inactive drug, and subsequently converts it into an active compound (see for example, U.S. Pat. No. 5,219,732). Such a construct would be capable of inducing the formation of a toxic substance in situ. For example, a catalytic sFv' molecule having  $\beta$ -lactamase-like activity can be designed to bind and catalyze the conversion of an inactive lactam derivative of doxorubicin into the active, cytotoxic form. Here the bispecific dimer, having binding affinities for both the preselected antigen and the cytotoxic-lactam derivative, is administered to an individual and allowed to accumulate at the desired location. The inactive, nontoxic cytotoxin-lactam derivative then is administered to the individual. When the derivative is complexed with the bispecific (sFv')<sub>2</sub> heterodimer in situ the active form of the drug is released, enhancing both the cytotoxicity and specificity of the drug.

#### Hybrid sFv' Considerations.

In a preferred aspect of the invention a humanized single-chain Fv is envisioned whereby the recombinant sFv' contains CDRs of the murine 741F8 antibody interposed between human FR sequences to generate a humanized c-erbB-2 binding protein. The humanized Fv would be capable of binding c-erbB-2 while eliciting little or no immune response when administered to a patient. A nucleic acid sequence encoding a humanized sFv may be designed and constructed as follows.

FR regions identified by homology searches of the GenBank database can be introduced into an sFv of interest by site-directed mutagenesis to reproduce the corresponding human sequence. Alternatively, homologous human V<sub>H</sub> and V<sub>L</sub> sequences can be derived from a collection of PCR-cloned human V regions, after which the human FR sequences can be ligated with murine CDR regions to create humanized V<sub>L</sub> and V<sub>H</sub> genes. A humanized sFv hybrid thus can be created, for instance, where the human FR regions of the human myeloma antibody are introduced between the murine CDR sequences of the murine monoclonal antibody 741F8. The resulting sFv, containing the sequences FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4, contains a murine binding site in a human framework.

By directly sequencing the DNA or RNA in a hybridoma secreting an antibody to a preselected antigen, or by obtaining the sequence from the literature, one skilled in the art can essentially produce any desired CDR and FR sequence. Expressed sequences subsequently may be tested for binding and empirically refined by exchanging selected 5 amino acids in relatively conserved regions, based on observations of trends of amino acid sequences in data bases and/or by using computer-assisted modeling techniques. Significant flexibility in  $V_H$  and  $V_L$  design is possible because alterations in amino acid sequences may be made at the DNA level.

Of course, the processes for manipulating, amplifying, and recombining DNAs that 10 encode amino acid sequences of interest are generally well known in the art (see, for example, Sambrook et al., 1989, Molecular Cloning A Laboratory Manual, 2nd ed. Cold Spring Harbor Press), and therefore, are not described in detail herein. Similarly, methods for identifying the isolated V genes encoding antibody Fv regions of interest are well understood and are described in the patent and other literature.

15

#### Expression of Recombinant sFv Proteins.

The resulting sFv DNA constructs then are integrated into expression vectors and transfected into appropriate host cells for protein expression. After being translated, the protein may be purified from the cells themselves or recovered from the culture medium.

20 The expression vectors also may include various sequences to promote correct expression of the recombinant protein. Typical sequences include transcription promoters and termination sequences, enhancer sequences, preferred ribosome binding site sequences, preferred mRNA leader sequences, preferred protein processing sequences, preferred signal sequences for protein secretion, and the like. The DNA sequence encoding the gene of interest also may be manipulated to remove potentially inhibiting sequences or to minimize unwanted secondary structure formation. The resulting synthetic genes can be expressed in appropriate prokaryotic hosts such as various strains 25 of E. coli, or in eucaryotic hosts such as Chinese hamster ovary cells (CHO), mouse myeloma, hybridoma, transfectoma, and human myeloma cells. The currently preferred expression system for the present invention is E. coli, as disclosed herein.

30 When the gene is to be expressed in E. coli, it is cloned into an expression vector downstream of a strong promoter sequence, such as Trp or Tac, and optionally also may include a gene coding for a leader polypeptide, such as the fragment B (FB) of staphylococcal protein A. The resulting fusion protein, when expressed, accumulates in

retractile bodies (also known as inclusion bodies) in the cytoplasm, and may be harvested after disruption of the cells by French press or sonication. The proteins then are solubilized, and refolded *in vitro*, as described herein. Where the construct is engineered as a fusion protein, the protein is solubilized and the leader sequence preferably cleaved before renaturation. The cleavage site for the leader sequence preferably is immediately adjacent to the sFv polypeptide chain and includes one amino acid or a sequence of amino acids exclusive of any one amino acid or amino acid sequence found in the amino acid structure of the single polypeptide chain.

The cleavage site preferably is designed for specific cleavage by a selected agent. Endopeptidases are preferred, although non-enzymatic (e.g., chemical) cleavage agents may be used. Many useful cleavage agents, for instance, cyanogen bromide (CNBr), dilute acid, trypsin, *Staphylococcus aureus* V-8 protease, post-proline cleaving enzyme, blood coagulation Factor Xa, enterokinase, and renin, recognize and preferentially or exclusively cleave at particular cleavage sites. One currently preferred peptide sequence cleavage agent is V-8 protease. The currently preferred cleavage site is at a Glu residue. Other useful enzymes recognize multiple residues as a cleavage site, e.g., factor Xa (Ile-Glu-Gly-Arg) or enterokinase (Asp-Asp-Asp-Asp-Lys). Dilute acid preferentially cleaves the peptide bond between Asp-Pro residues, and CNBr in acid cleaves after Met, unless it is followed by Tyr.

Alternatively, the engineered gene may be incorporated into a vector without a sequence encoding a leader polypeptide, and the engineered gene expressed to produce a polypeptide chain that is secreted into the *E. coli* periplasmic space. The secreted protein then can be isolated and, optionally, purified further using standard methodologies. (See, for example, Pack et al. (1992) *Biochem* 31:1579-1584.)

If the engineered gene is to be expressed in eucaryotic hybridoma cells, the conventional expression host for immunoglobulins, the gene preferably is inserted into an expression vector containing, for example, the immunoglobulin promoter, a secretion signal, and immunoglobulin enhancers. This plasmid also may contain sequences encoding other polypeptide chains, including part or all of a toxin, enzyme, cytokine, or hormone. The gene then is transfected into myeloma cells via established electroporation or protoplast fusion methods. The transfected cells then may express V<sub>H</sub>-linker-V<sub>L</sub>-tail or V<sub>L</sub>-linker-V<sub>H</sub>-tail single-chain Fv' polypeptide chains.

The sFv polypeptide chains can be expressed as either inactive or active polypeptide chains. Spontaneously refolded sFv polypeptide chains can be obtained from

either prokaryotic or eukaryotic expression systems when the polypeptide chains are secreted for instance, either into the E. coli periplasmic space or the mammalian cell culture medium. These spontaneously refolded polypeptide chains readily can be purified by affinity chromatography. Where the sFv polypeptide chains are obtained in an 5 unfolded, inactive sFv form, for instance, when overexpression of the sFv polypeptide chain in E. coli results in the formation of inclusion bodies, the proteins can be refolded in vitro. Briefly, inclusion bodies are harvested by centrifugation, the sFv, solubilized with denaturants such as guanidine hydrochloride (GuHCl) or urea, and then refolded by dilution of the denaturant under appropriate redox (reduction/oxidation) conditions (see 10 below). The refolded sFv polypeptide chains then can be purified by affinity chromatography. Details for the isolation of inclusion bodies, solubilization and renaturation of the sFv polypeptide chains are well known in the art (see for example, U.S. Pat. No. 5,091,513 and Huston et al., 1988, *supra*).

15 Dimerization and Purification of the sFv Polypeptides.

The sFv' monomers of the present invention can be dimerized in vivo or in vitro. In the in vivo approach, two sFv' genes can be cotransfected into the host cell wherein the coexpressed sFv' polypeptide chains spontaneously dimerize. Alternatively, the refolded, secreted sFv' polypeptide chain monomers can be isolated from two expression hosts and 20 subsequently dimerized in vitro.

In a preferred aspect of the invention, the sFv' polypeptide chains comprising a single cysteine C-terminal tail residue are expressed in E. coli and form inclusion bodies. The resulting sFv' polypeptide chains are solubilized with denaturants and renatured in vitro, either in the presence or absence of exogenously added glutathione. Surprisingly, 25 the additional C-terminal cysteine residues apparently do not interfere with the refolding process. In some cases however, sFv and sFv' constructs may refold poorly in vitro. These constructs can be "preoxidized prior" to refolding as taught in Huston et al., (1991) Meth. Enzymol. 203:46-88, or, alternatively, the polypeptide chains can be secreted across a membrane bilayer. The latter process spontaneously separates the extra C-terminal 30 cysteine residue from the cysteine residues normally found in the Fv domain, minimizing inappropriate disulfide bond formation. Secretion is the preferred method if the sFv' constructs refold poorly in vitro.

Following renaturation of the sFv' monomers, (sFv')<sub>2</sub> dimers readily can be prepared in vitro by air oxidation if cysteine amino acids are present in the C-terminal tail

sequences. Alternatively, sulphydryl specific crosslinking reagents, for instance, the BMH crosslinker or the MCA-peptide-MCA bridge may be used to covalently couple two sFv' chains. The resultant homo or heterodimers, then can be purified by standard size exclusion chromatography. However, when (sFv)<sub>2</sub> heterodimers are required, then a 5 preferred purification protocol uses a sequential two step affinity chromatography procedure. Briefly, the heterodimer is exposed to a first chromatographic system having an epitope that interacts specifically with one sFv of the heterodimer. The eluant containing the heterodimer is then exposed to a second system having an epitope that interacts specifically with the other sFv. For specific details of the dimerization and 10 purification procedures, see Example 2.

Considerations for In Vivo Administration.

The dimeric constructs may be administered either by intravenous or intramuscular injection. Effective dosages for the single-chain Fv constructs in antitumor therapies or in 15 effective tumor imaging can be determined by routine experimentation, keeping in mind the objective of the treatment.

The pharmaceutical forms suitable for injection include sterile aqueous solutions or dispersions. In all cases, the form must be sterile and must be fluid so as to be easily administered by syringe. It must be stable under the conditions of manufacture and 20 storage, and must be preserved against the contaminating action of microorganisms. This may, for example, be achieved by filtration through a sterile 0.22 micron filter and/or lyophilization followed by sterilization with a gamma ray source.

Sterile injectable solutions are prepared by incorporating the desirable amount of the constructs, disclosed herein, into an appropriate solvent, such as sodium phosphate-buffered saline (PBS), followed by filter sterilization. As used herein, "a physiologically 25 acceptable carrier" includes any and all solvents, dispersion media, antibacterial and antifungal agents that are non-toxic to humans, and the like. The use of such media and agents as pharmaceutically active substances are well known in the art. The media or agent must be compatible with maintenance of proper conformation of the single-chain Fv polypeptide chains, and its use in the therapeutic compositions. Supplementary active 30 ingredients can also be incorporated into the compositions.

A preferred remotely detectable moiety for in vivo imaging includes the radioactive atom Technetium<sup>99m</sup> (<sup>99m</sup>Tc), a gamma emitter with a half-life of about 6 hours. Non-radioactive moieties also useful in imaging include nitroxide spin labels as

well as lanthanide and transition metal ions all of which induce proton relaxation in situ. In addition to immunoimaging, the complexed radioactive moieties may be used in standard radioimmunotherapy protocols to destroy the targeted cell. Preferred nucleotides for high dose radioimmunotherapy include the radioactive atoms <sup>90</sup>Yttrium (<sup>90</sup>Yt),  
5 <sup>131</sup>Iodine (<sup>131</sup>I) and <sup>111</sup>Indium (<sup>111</sup>In).

Either the single polypeptide chain sFv' itself, or the spacer groups for linking the sFv' constructs can be labeled with radioisotopes such as <sup>131</sup>I, <sup>111</sup>In and <sup>99m</sup>Tc. <sup>99m</sup>Tc and <sup>111</sup>In are preferred because they can be detected with gamma cameras and have favorable half-lives for in vivo imaging applications. The single polypeptide chains can be labeled, 10 for example, with radioactive atoms such as <sup>90</sup>Ty, <sup>99m</sup>Tc or <sup>111</sup>I via a conjugated metal chelator (see, e.g., Khaw et al., 1980, Science 209: 295; U.S. Pat. No. 4,472,509; U.S. Pat. No. 4,479,930), or by other standard means of linking isotopes to proteins, known to those with skill in the art (see for example, Thankur et al., 1991, J. Immunol. Methods 237: 217-224).

15

The invention is illustrated by the following Examples, which are not intended to be limiting in any way.

20

## EXAMPLES

### Example 1. Synthesis and Expression of the sFv Constructs (741F8, 26-10 and 520C9).

The construction of several sFv genes using different but standard recombinant DNA technology, well known to those having ordinary skill in the art, is described below. These procedures include the amplification of the V<sub>H</sub> and V<sub>L</sub> gene sequences by PCR, the 25 ligation of appropriate synthetic DNA duplexes and the cloning of V<sub>H</sub> or V<sub>L</sub> genes by colony hybridization.

#### A. 741F8 sFv'.

The V<sub>H</sub> and V<sub>L</sub> genes of the 741F8 anti-c-erbB-2 monoclonal antibody were isolated from the cDNA of the parental 741F8 hybridoma line by PCR using primers 30 homologous to the N-terminal coding regions of V<sub>H</sub>, V<sub>L</sub>, C<sub>H1</sub>, and C<sub>L</sub>. The PCR-amplified V<sub>H</sub> and V<sub>L</sub> genes were isolated by polyacrylamide gel electrophoresis and cloned into a pUC cloning vector. The first FR region of the 741F8 V<sub>H</sub> gene however contained spurious mutations due to the PCR procedure. Errors were rectified by the replacement of the first

70 nucleotides of 741F8  $V_H$  with a similar sequence from 520C9  $V_H$ , another c-erbB-2 specific monoclonal antibody.

5 Restriction sites then were introduced into the ends of the heavy and light chain variable gene segments by site-directed mutagenesis (Kunkel *et al.*, 1985, *Proc. Natl. Acad. Sci. USA* **82**: 488-492). A Nco I site encoding methionine was positioned at the N-terminus of  $V_H$  for expression in *E. coli*. A Sac I site was created at the 3' end of  $V_H$  gene. A Xho I site, together with an adjacent Eco RV site, were created at the N-terminus of  $V_L$ . A stop codon and a Pst I site were placed at the C-terminal end of  $V_L$ .

10 The single-chain Fv gene was constructed by connecting the  $V_H$  and  $V_L$  genes together with a DNA sequence encoding the 14 residue polypeptide linker, (Ser<sub>4</sub>Gly)<sub>2</sub>Ser<sub>4</sub>, as set forth as amino acids 122 through 135 in the Sequence Listing as SEQ. ID. NOS.: 1 and 2.

15 A synthetic DNA duplex encoding the C-terminal amino acid sequence, (Gly)<sub>4</sub>-Cys was inserted into a Hpa I site located near the stop codon at the 3' end of the 741F8 sFv gene. The resulting 741F8 anti-c-erbB-2 sFv' gene was excised from the pUC cloning vector, with the restriction enzymes Nco I and Bam HI (a Bam HI site is located 3' to the C-terminal Pst I site), and inserted into the same sites of a commercial T7 expression vector pET-3d (In-vitrogen, Inc.). The resulting gene, set forth in the Sequence Listing as SEQ. ID. NOS.: 1 and 2, was transformed into *E. coli* BL21-DE (In-vitrogen, Inc.).

20 Protein expression was induced by the addition of IPTG to the culture medium.

#### B. 26-10 sFv'

25 Construction of the anti-digoxin 26-10 sFv has been described previously (Huston *et al.*, 1988, *Proc. Natl. Acad. Sci. USA* **85**; 5879-5883, and U.S. Pat. No. 5,091,513, both of which are hereby incorporated by reference). Briefly, the synthetic gene was constructed by ligating multiple synthetic DNA duplexes together. The C-terminal DNA duplex coding for the amino acid sequence (Gly)<sub>4</sub>-Cys subsequently was ligated into a Hpa I restriction site close to the 3' end of the 26-10 sFv gene. The resulting sFv' gene, set forth in the Sequence Listing as SEQ. ID. NOS.: 3 and 4, was then inserted into the *E. coli* expression vector pET-3d. This plasmid was subsequently transformed into *E. coli* BL21-DE (In-vitrogen, Inc.) and protein expression induced by the addition of IPTG to the culture medium.

#### C. 520C9 sFv.

The 520C9 sFv was generated by linking together the  $V_H$  and  $V_L$  genes, cloned from a 520C9 hybridoma cDNA library, with a serine rich linker. Briefly, the  $V_H$  and  $V_L$  genes were cloned from the 520C9 hybridoma cDNA library using probes directed toward the antibody constant (C) and joining (J) regions. Appropriate restriction sites were 5 introduced at the ends of each gene by site-directed mutagenesis (Kunkel *et al.*, 1985, *Proc. Natl. Acad. Sci. USA* 82: 488-492). The  $V_H$  and  $V_L$  genes were then ligated together with a serine rich linker. The resulting 520C9 sFv gene, set forth in the Sequence Listing as SEQ. ID. NOS.: 5 and 6, was transformed into the *E. coli* expression vector and expressed as described above and in co-pending U.S.S.N. 831,967, incorporated therein by 10 reference.

Example 2. Renaturation, Dimerization and Purification of sFv Proteins.

A. Renaturation and Purification of sFv monomers.

Protocols for renaturing sFv monomers derived from *E. coli* inclusion bodies are 15 described below. In separate experiments the 7418, 26-10 and 520C9 sFv polypeptides were expressed in *E. coli*. The unfolded sFv proteins were solubilized from inclusion bodies and refolded under appropriate redox conditions. The refolded sFv polypeptide chains were purified by affinity chromatography or by a combination of ion-exchange and size exclusion chromatography when affinity chromatography was not feasible or 20 expedient.

Renaturation of 741F8 sFv'.

Inclusion bodies containing the 741F8 sFv' proteins were washed in a buffer containing 25 mM Tris, 10 mM EDTA, 1.5M GuHCl, pH 8.0 and solubilized in 25 mM 25 Tris, 10 mM EDTA, 7M GuHCl, pH 9.0 to an  $OD_{280\text{ nm}}$  of about 25-50. The sample was reduced overnight at room temperature by the addition of dithiothreitol (DTT) to a final concentration of 10 mM. The thiol groups were converted into mixed disulfides with glutathione by the addition of solid oxidized glutathione to a final concentration of 100 mM. The solution was adjusted to pH 9.0 and incubated for 4 hr at room temperature. 30 The 741F8 sFv' polypeptide chains then were refolded *in vitro* to generate stable monomers with their C-terminal cysteines remaining blocked with glutathione. The 741F8 sFv' mixed disulfide preparation was diluted to an  $OD_{280}$  of about 0.15 by the addition of 10 mM Tris, 4 mM EDTA, 6M urea, pH 8.5 at 4° C. After two hours an equal volume of 10 mM Tris, 4 mM EDTA, 1 mM reduced glutathione, pH 8.5, precooled to

4°C., was added with rapid mixing to reduce the urea concentration to 3M. After dilution, the samples were allowed to renature for 72 hr at 4° C.

Renaturation of 26-10 sFv'.

5 Inclusion bodies containing the 26-10 sFv' proteins were washed with 25 mM Tris, 10 mM EDTA and solubilized in 6M GuHCl, 25 mM Tris, 10 mM EDTA, pH 8.7 to an OD<sub>280 nm</sub> of about 10 to 20. The dissolved proteins were reduced by overnight incubation at room temperature after the addition of DTT to 10 mM. The reduced protein could also be blocked with oxidized glutathione as noted above for the 741F8 sFv' polypeptide. The 10 reduced, denatured 26-10 sFv' polypeptides were refolded in a manner similar to that for the 741F8 sFv' by diluting the preparation into a buffer containing 3M urea, 0.1 mM oxidized and 0.01 mM reduced glutathione to give a final protein concentration of about 0.15 mg/ml. After overnight incubation at 4° C., the mixture was dialyzed against PBS containing 0.05M KH<sub>2</sub>PO<sub>4</sub>, 0.15M NaCl, pH 7 for two days at 4° C.

15

Renaturation of 520C9 sFv.

20 The inclusion bodies containing the 520C9 sFv were washed with 25 mM Tris, 10 mM EDTA, pH 8.0, 1M GuHCl and solubilized in 25 mM Tris, 10 mM EDTA, 6M GuHCl, 10 mM dithiothreitol (DTT), pH 9.0. The material was ethanol precipitated and resuspended in 25 mM Tris, 10 mM EDTA, 6M urea, 10 mM DTT, pH 8.0 and fractionated by ion exchange chromatography to remove contaminating nucleic acids and *E. coli* proteins before renaturation of the sFv. The material that did not bind to a DEAE Sepharose Fast Flow (FF) column was precipitated by lowering the pH to 5.5 with 1M acetic acid. The pellet was resolubilized in 25 mM Tris, 10 mM EDTA, 6M GuHCl, 10 mM DTT, pH 9.0 and oxidized by overnight incubation at room temperature following dilution into a buffer containing 25 mM Tris, 10 mM EDTA 6M GuHCl, 1 mM oxidized glutathione, 0.1 mM reduced glutathione, pH 9.0. After overnight oxidation the sample was dialyzed against 10 mM NaH<sub>2</sub>PO<sub>4</sub>, 1 mM EDTA, 150 mM NaCl, 500 mM urea, pH 8.0 and the sample clarified by filtration through a membrane with a 100 kD mol. wt. cut-off prior to purification on a c-erbB-2 affinity column.

Purification of the refolded sFv Polypeptides.

30 The refolded 26-10 sFv' polypeptide chains were purified by ouabain-Sepharose affinity chromatography, as described for the 26-10 sFv constructs (Huston, *et. al.*, 1988,

5 Proc. Natl Acad. Sci. USA **85**; 5879-5883 and Tai, *et. al.*, 1990, Biochem. **29**, 8024-3080, both of which are hereby incorporated by reference). The refolded 520C9 sFv polypeptide chain was similarly purified using a c-erbB-2-agarose affinity column. In this case, the refolded samples were loaded onto a c-erbB-2 affinity column, the column washed with PBS, and the 520C9 sFv polypeptides eluted with PBS pH 6.1 containing 3M LiCl. The buffer was then exchanged by dialysis. The c-erbB-2 affinity column preferably was prepared by linking the extracellular domain of c-erbB-2 onto agarose beads.

10 Briefly, the c-erbB-2 sequence coding for its extracellular domain (ECD) was derived from the baculovirus expression vector described previously (Ring *et al.*, 1992, Mol. Immunol. **28**; 915-917). A DNA duplex encoding the His<sub>6</sub> peptide was ligated to the 3' end of the ECD gene, and the construct expressed in CHO cells. The ECD polypeptide was purified from the CHO cell culture medium on an IMAC metal affinity column (Pharmacia, Piscataway, N.J.), as described in Skerra, *et al.*, 1991, Bio/Technology **9**: 273-278, and the eluted ECD proteins attached onto agarose beads to generate the c-erbB-2-agarose affinity resin.

15 The renatured 741F8 sFv' polypeptides were purified by a combination of ion exchange and size exclusion chromatography. Briefly, the renatured 741F8 sFv' preparation was passed through a DEAE-cellulose column and the 741F8 sFv' in the unbound fraction adjusted to pH 5.0 before loading on an S-Sepharose FF column. The 20 741F8 sFv' polypeptide chains were eluted with PBS containing 2 mM EDTA and 3M urea, and dialyzed against 10 mM Tris, 2 mM EDTA, 20 mM NaCl, pH 7.5 at 20° C. The precipitate was harvested by centrifugation, dissolved in a suitable buffer, and passed through a Q-Sepharose FF column. The unbound material was adjusted to pH 5.5 and reloaded onto a S-Sepharose FF column. The 741F8 sFv' polypeptides were eluted with a 25 PBS, 2 mM EDTA, 100 mM NaCl, 3M urea buffer and dialyzed against PBS, 2 mM EDTA. The precipitate was harvested again by centrifugation, dissolved in a suitable buffer, sucrose added to 5% (w/v), and the 741F8 sFv' concentrated to 5 mg/ml in a YM10 membrane concentrator (Amicon). The 741F8 sFv' polypeptide chains were fractionated by gel filtration chromatography using a S-200 HR column (Pharmacia LKB Biotechnology) and a PBS, 2 mM EDTA buffer.

#### B. Dimerization of the sFv' constructs

Dimerization of sFv' monomers can be induced using standard crosslinking conditions. Where disulfide bond formation is desired, the monovalent sFv' polypeptide

chains initially are deblocked by mild reduction and (sFv')<sub>2</sub> dimers formed by crosslinking the sFv' polypeptides either by disulfide linkages or by thioether linkages with the BMH or MCA-peptide-MCA crosslinking reagents.

In order to generate disulfide linked constructs the purified 741F8 and 26-10 sFv' preparations were dialyzed against 50 mM Tris, 150 mM NaCl, pH 8.5. The C-terminal glutathionyl blocking groups were removed by the addition DTT to a concentration of 2 mM followed by overnight incubation at room temperature. Excess reducing agent was removed by extensive dialysis against 50 mM Tris, 150 mM NaCl, pH 8.5, during which the majority of the sFv' polypeptides oxidized into the homodimeric form.

In order to generate BMH and MCA-peptide-MCA crosslinked constructs, sFv' polypeptide chains in PBS first were reduced for two hours at room temperature by the addition of DTT to a final concentration of 1 mM. The samples were desalted by gel filtration chromatography using a PBS, 1 mM EDTA buffer. A 4-5 fold molar excess of either the BMH or MCA-peptide-MCA linkers, both dissolved in dimethylsulfoxide, were added to the reduced protein and incubated for at least 12 hours at room temperature. The resulting dimers were then purified by HPLC gel filtration chromatography.

A modification of the procedure of Brennan, *et al.* (1985, *Science* 229: 81-83) may be used to generate disulfide linked sFv' heterodimers. For example, in order to link the 741F8 and 26-10 sFv' polypeptides a thionitrobenzoate (TNB) derivative of the 26-10 sFv' (26-10 sFv'-TNB) was mixed with mildly reduced 741F8 sFv'. The 26-10 sFv'-TNB was prepared by reducing the 26-10 sFv' in PBS with 15 mM 2-mercaptoethylamine for 30 minutes at room temperature. The reducing agent was removed by gel filtration and the reduced 26-10 sFv' reacted with 2.2 mM dithionitrobenzoate (DTNB) for 3 hours. The active 26-10 sFv'-TNB was adsorbed onto onto ouabain-Sepharose. The glutathionyl blocked 741F8 sFv' monomer in 25 mM Tris, 150 mM NaCl, pH 8.2 was reduced for 2 hours at room temperature by the addition of DTT to a final concentration of 1 mM. The excess DTT was removed by gel filtration and the reduced 741F8 sFv' reacted overnight at room temperature with the 26-10 sFv'-TNB complexed to ouabain-Sepharose. The progress of the reaction was monitored spectroscopically at 412 nm, the absorbance maximum of the TNB anion.

### C. Purification of (sFv')<sub>2</sub> Dimers.

The (sFv')<sub>2</sub> homodimers may be separated from the sFv' monomers by gel filtration chromatography. Following dimerization, the sFv' preparations are dialyzed against PBS

containing 1 mM EDTA, 3M urea, 0.03% azide, to disrupt any non-covalent homodimers and fractionated by HPLC on a TSK-G20000SW column using the same buffer. The procedure requires two passes for purification of the (sFv')<sub>2</sub> homodimers to homogeneity. The purified homodimers may be dialyzed either against PBS or any other suitable buffer prior to use.

The (sFv')<sub>2</sub> heterodimers can be separated by a two step affinity chromatography procedure taking advantage of the bivalent nature of the dimer. For instance, during the purification of the 741F8/26-10 heterodimer the mixture initially was loaded onto an ouabain-Sepharose column, washed with a PBS, 1M NaCl buffer, to remove any non-specifically adsorbed material, and rewashed with PBS to reduce the salt concentration. The reactive 26-10 sFv' species bound to the resin were eluted with 20 mM ouabain in PBS and the eluate dialyzed against PBS to remove the cardiac glycoside. The 741F8/26-10 heterodimers were then repurified on a c-erbB-2-agarose affinity column taking advantage of the ECD binding site in the heterodimer. After the preparation was loaded onto the c-erbB-2 affinity column, it is washed with PBS and the (sFv')<sub>2</sub> heterodimer eluted with 25 mM Tris, 10 mM EDTA, 5M LiCl, pH 6.8. Prior to use, the buffer was exchanged with PBS by dialysis.

### Example 3. Immunoreactivity of the Monomeric and Dimeric sFv Polypeptides.

#### A. Radiolabeling of the sFv' Constructs.

The sFv' polypeptides may be labeled by the chloramine-T method as described (DeNardo, *et al.*, 1986, *Nucl. Med. Biol.* **13**: 303-310). Briefly, 1.0-2.0 mg of sFv' was combined with <sup>125</sup>I [14-17 mCi/ $\mu$ g] (Amersham, Arlington Heights, Ill.) at an iodine to protein ratio of 1:10 in a 12 x 75 mm plastic test tube. 10  $\mu$ l [1 mg/ml] of chloramine-T (Sigma, St. Louis, Mo.) per 100  $\mu$ g of protein was added and the mixture incubated for three minutes at room temperature. After the reaction was terminated, unincorporated <sup>125</sup>I was separated from the labeled sFv' by the spun-column method of Meares, *et al.*, 1984, *Anal. Biochem.* **142**: 68-78. Specific activities of 0.2-1.0 mCi/mg for the <sup>125</sup>I-labeled products may be routinely obtained.

#### B. Competition ELISA

In order to prepare c-erbB-2, SK-Br-3 breast cancer cells (Ring *et al.*, 1989, *Cancer Res.* **49**: 3070-3080), were harvested and resuspended in 10 mM NaCl, 0.5% Nonidet-P40,

pH 8. Insoluble debris was removed by centrifugation and the extract filtered through 0.45 Millex HA and 0.2 Millex GV filters. 40  $\mu$ l of the extract was added to each well of a 96 well plate and incubated overnight at 37° C. The plates then were washed with PBS and non-specific binding sites blocked following the addition of PBS containing 1% skim milk by incubation for one hour at room temperature. The sFv and 520C9 Fab samples, 5 diluted in PBS, were added to the wells and incubated for 30 mins at room temperature. A control containing only dilution buffer was also included.

In order to quantitate the reaction, 20  $\mu$ l of a 520C9-horseradish peroxidase (HRP) 10 probe (Zymed Labs., South San Francisco, Calif.), diluted to 14  $\mu$ l/ml in PBS containing 1% skim milk, was added to each well and incubated for one hour at room temperature. The plate was then washed four times with PBS, the peroxidase substrate added and incubated for 30 minutes at room temperature. The reaction was quenched with H<sub>2</sub>SO<sub>4</sub> and the OD<sub>150nm</sub> values measured.

15 FIG. 3 compares the binding ability of the parental 520C9 Fab fragment, together with the 520C9 sFv single-chain binding protein. The 520C9 sFv samples included the material obtained following renaturation of the polypeptide *in vitro*, a sample purified on a c-erbB-2 agarose affinity column, and the material that did not bind to the column. The fully purified 520C9 sFv polypeptide exhibits an affinity for c-erbB-2 indistinguishable from the parent 520C9 Fab fragment.

20

### C. Biodistribution Studies.

25 *In vivo* immunotargeting tissue imaging studies were performed using standard procedures. Approximately 2.5 x 10<sup>6</sup> SK-OV-3 cells (a human ovarian cancer cell line that expresses c-erbB-2 on the cell surface) in log phase were implanted subcutaneously onto the hips of four to six week old C.B17/ICI-scid mice. Three days after Lugol's solution was placed in the drinking water to block the accumulation of radioiodine in the thyroid, the mice were used in the biodistribution assays.

The radiolabeled sFv' and Fab preparations were diluted in PBS for these studies. The biodistribution of the glutathionyl-blocked 741F8 sFv' monomers, and the 741F8 and 30 26-10 (sFv')<sub>2</sub> constructs were compared after identical doses of the radiolabeled protein was administered by injection in each case. The total injected doses were determined by counting each animal on a Series 30 multichannel analyzer/probe system (probe model #2007, Canaberra, Meridian, Conn.). Groups of 3-6 mice were sacrificed twenty four

hours after injection, the tumors and organs were removed, weighed and counted in a gamma counter to determine the amount of radiolabel incorporated into the tissues. From these measurements, the percentage of the initial injected dose incorporated per gram of tissue (%ID/gram) or the amount of label incorporated into the tumor relative to the amount of radiolabel incorporated into the other organs (T:O ratio) were determined. For specific details see DeNardo, *et al.*, 1977, *Cancer*, **40**: 2923-2929, or Adams, *et al.*, 1992, *Antibody, Immunoconjugates, and Radiopharmaceuticals* **5**: 81-95, both of which are hereby incorporated by reference. Specificity indices also can be determined by dividing the T:O ratios of the  $^{125}\text{I}$ -741F8 sFv' by the corresponding T:O ratios of the  $^{125}\text{I}$ -26-10 sFv'.  
5 The results of the biodistribution studies 24 hours post administration are summarized in Figs. 4 and 5. The mean standard error (SEM) for each value is less than 30%, except where indicated.  
10

The disulfide linked 741F8 (sFv')<sub>2</sub> homodimers exhibit identical tumor specificities when compared to the monomeric 741F8 sFv' polypeptide chains. The T:O ratios of the 15 741F8 sFv' constructs consistently exceed those for the 26-10 sFv' constructs, demonstrating the binding specificity of the 741F8 constructs for the tumors (FIG. 4). In addition, the 741F8 (sFv')<sub>2</sub> dimers generally exhibit higher T:O ratios relative to that of the monomeric species, particularly for the disulfide bonded sFv' 741F8 (sFv'-(Gly)<sub>4</sub>Cys)<sub>2</sub> and the MCA linked 741F8 (sFv')<sub>2</sub> homodimers. In addition, the 741F8 (sFv')<sub>2</sub> homodimers 20 localize in greater amounts in the tumors relative to the monomeric sFv' species (FIG. 5).

In a separate comparative study with  $^{125}\text{I}$ -labeled 26-10 (sFv')<sub>2</sub> and the following species of  $^{125}\text{I}$ -labeled 741F8: sFv' monomers, Fab, disulfide linked (sFv'-Gly<sub>4</sub>Cys)<sub>2</sub> homodimers, and MCA- and BMH-linked (sFv')<sub>2</sub> homodimers, the *in vivo* tumor localization properties of these molecules were compared (% ID/gram tumor tissue, see 25 FIG. 6). As is evident from the figure, the tumor localization properties of all of the dimeric 741F8 (sFv')<sub>2</sub> constructs are significantly greater than those observed with the 741F8 Fab, the 741F8 sFv' monomer and the 26-10 (sFv')<sub>2</sub> dimer (FIG. 6). The results demonstrate that the increased apparent avidity and enhanced *in vivo* imaging of the (sFv')<sub>2</sub> dimer is due, at least in part, to its improved retention in tumor tissue.

30

All publications, patents and patent applications are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent or

patent application was specifically and individually indicated to be incorporated by reference in its entirety.

5

### *Equivalents*

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

10